

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 18:54:32 ; Search time 4770 Seconds  
(without alignment)  
10613.308 Million cell updates/sec

Title: US-10-630-636-1  
Perfect score: 1330  
Sequence: 1 gtgcgagatccaccaccga.....gacccaaaaaaaaaaaaa 1330

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc1:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_g881:\*  
9: gb\_g882:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	447	33.6	478	7	CF326660 JMT1--06-
2	432.2	32.5	461	9	CL794965 OR CBA000
3	370.2	27.8	689	5	BQ483926 WHE514_C
4	352.6	26.5	359	9	CL728595 OR BBA005
5	311.2	23.4	815	6	CA185031 SCBSST309
6	303.4	22.8	618	9	CL782098 OR BBA009
7	274.4	20.6	884	9	CG527740 OGXCQ48TH
8	254.6	19.9	689	9	CG710311 OGDH26TC
9	252.2	19.7	804	9	CG366578 OGAM96TV
10	252.2	19.7	853	9	CG362076 OGICL41TH
11	252.2	19.7	933	9	CG319911 OG1AD74TH
12	250.6	19.6	807	9	CL968676 OaIFCC017
13	254.8	19.2	740	9	CG751932 ZMWBBD013
14	253.6	19.1	468	4	BM038003 U001D12 O
15	251.2	18.9	653	4	CC752461 ZMWBBD013
16	242.4	18.2	560	9	CL620317 OR BBA001
17	240.2	18.1	921	9	CG257748 OGXCQ48TV
18	233.8	17.6	759	8	BZ525067 OGAT80TC
19	224.4	16.9	606	6	CA235113 SCCCF1A09
20	223	16.8	874	9	CG319919 OG1AD74TH
21	223	16.8	912	9	CG366864 OGAM96TV
22	220.6	16.6	914	9	CG362086 OGICL41TV
23	218.2	16.4	515	9	CL707638 OR BBA002
24	212.8	16.0	763	9	CC693661 OGFAU36TV

25	204.4	15.4	421	8	BH874076 hp48c11.9
26	204.4	15.4	620	6	CA273712 SCR1SD100
27	199	15.0	921	9	CL297228 ZMWBBD000
28	198.8	14.9	503	6	CA151962 SCJPR2201
29	198.2	14.9	661	4	B1954597 HVSMEM001
30	195.2	14.7	705	6	CA243768 SCQSF1407
31	191.6	14.4	454	6	CA692850 wlm96_PkO
32	190.8	14.3	643	6	CA233892 SCSCFLA0C
33	186.6	14.0	639	7	CF428085 PH1_12 GO
34	185.6	14.0	628	2	BE356185 DG1_123 F
35	185	13.9	348	4	BM500306 PACQ00000
36	180.2	13.5	477	4	BM376711 EBem05_SQ
37	180.2	13.5	666	4	BM072319 MEST64-C1
38	180.2	13.5	725	4	BM075490 MEST355-H
39	179.6	13.5	900	9	CL960237 OaIFCC004
40	178.6	13.4	601	4	BM333529 MEST157-A
41	175.8	13.2	663	6	CA292935 SCR1FL802
42	175.2	13.1	517	9	CC710318 OGDH26TC
43	174.6	13.1	741	7	CO533588 3530_1_22
44	172.2	12.9	586	7	CN010050 WHE3865 H
45	171.8	12.9	690	7	CO524463 3530_1_16

#### ALIGNMENTS

RESULT 1  
LOCUS CF326660 478 bp mRNA linear EST 18-AUG-2003  
DEFINITION JMT1--06-K20.g1 AtUMT-overexpressing transgenic rice lambda phage cDNA library (JMT1) Oryza sativa (japonica cultivar-group) CDNA clone JMT1--06-K20, mRNA sequence.  
VERSION CF326660.1 GI:33801577  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 478)  
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
Large-scale Sequencing Analysis of Rice ESTs  
Unpublished (2003)  
CONTACT: Nahm B.H.  
Genomics and Genetics Institute, Greengene Biotech Inc., Division of Bioscience and Bioinformatics, Wonyoung University  
Yongin, Kyeonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.  
Location/Qualifiers  
1..478  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:39947"  
/clone="JMT1--06-K20"  
/tissue\_type="leaf"  
/dev\_stage="14 days after germination"  
/lab\_host="E.coli SOLR"  
/clone\_lib="AtUMT-overexpressing transgenic rice lambda phage cDNA library (JMT1)"  
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end with XhoI site. mRNA was prepared from Arabidopsis jasmonate Carboxyl methyltransferase overexpression line."

ORIGIN  
Query Match 33.6%; Score 447; DB 7; Length 478;  
Best Local Similarity 98.3%; Pred. No. 4.2e-87;

	Matches	462;	Conservative	0;	Mismatches	6;	Indels	2;	Gaps	1;
QY	475	CTTGCTGGGGCTTGACCAAGTTTCGGCAAGGGGACCTGGGGAGCATCTCGGCCCACTTTCGT	534							
Db	9	CTTCGTGGGGCTTGACCAAGTTTCGGCAAGGGGACCTGGGGAGCATCTCGGCCCACTTTCGT	68							
QY	535	CATCTCGCGAGCGGCCACGACAGGTGGCCAGGCCACGCGCAGAAAGTACTTCATCTCGGCTTCAA	594							
Db	69	CATCTCGCGAGCGGCCACGACAGGTGGCCAGGCCACGCGCAGAAAGTACTTCATCTCGGCTTCAA	128							
QY	595	CTTCATGAACCGCGACCGCCGCGCTCCAGCATCCACGACATCACCAAGCTTACCGGCGG	654							
Db	129	CTTCATGAACCGCGACCGCCGCGCTTCAGACATCCACGACATCAACAGCTTACCGGCGG	188							
QY	655	CGATCAGGATCGCCGCGACAGGGGCGCCCGATCATCCGAGCCACACGAGGCCACGGGCAACC	714							
Db	189	CGATCAGGATCGCCGCGACAGGGGCGCCCGATCATCCGAGCCACACGAGGCCACGGGCAACC	248							
QY	715	CGCGG--CGGCGGCGCTGGGCGCCGCGGAGCATGAAGACCAACACACCAACCAACCCGGGCG	772							
Db	249	CGCGGCGCGCGAGNCGCTGGGCGCGCGGAGCATGAAGACCAACCAACCAACCCGGGCG	308							
QY	773	GGCGGCGCGCGCGCCCATGCGCATGTACACAGCGCGCGCCCATGGGCGCAACCCGTCGCGGCG	832							
Db	309	GGCGGCGCGCGCGCCCATGCGCATGTACACAGCGCGCGCCCATGGGCGCAACCCGTCGCGGCG	368							
QY	833	CACATGATGCCCGCGCGCGCTCGGACACGCGCGGTGTTCCTCGCGCGGACACGCGCGCTAC	892							
Db	369	CACATGATGCCCGCGCGCGCTCGGACACGCGCGGTGTTCCTCGCGCGGACACGCGCGCTAC	428							
QY	893	GTCGCGCCCGTTCGCTACCCGCGCGCTCCGCGCAAGATGACCAACCATATACG	942							
Db	429	GTCGCGCCCGTTCGCTACCCGCGCGCAACCGCGCAAGATGACCAACCATATACG	478							

RESULT	2
LOCUS	CL794965/c
DEFINITION	CL794965                461 bp   DNA             linear    GSS  06-AUG-2004
ACCESSION	OR_CBA0004J23.f_OR_CBA Oryza rufipogon genomic clone OR_CBA0004J23
VERSION	5 , genomic survey sequence.
KEYWORDS	CL794965                CL794965                GI:51016981
SOURCE	GSS.
ORGANISM	Oryza rufipogon Oryza rufipogon Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehretaceae; Oryzaceae; Oryza. 1 (bases 1 to 461) Kim,H., Yu,Y., Waisolski,M., Yost,D., Scum,D., Rao,K., Luo,M., Jetty,R., Kudina,D., Miller,C., Hatfield,J., Soderlund,C. and Wing,R. OMP project Unpublished (2004)
TITLE	Contact: Rod A. Wing
JOURNAL	Arizona Genomics Institute
COMMENT	University of Arizona Forbes Building Room 303, Tucson, AZ 85721-0036, USA Tel.: 520 626 9595 Fax: 520 621 1259 Email: http://genome.arizona.edu
AUTHORS	PCR Primers FORWARD: TAA TAG GAC TCA CTA TAG GG BACKWARD: CAC TCA TTA GCC ACC CCA Plate: 0004 row: J column: 23 Seq primer: TAA TAG GAC TCA CTA TAG GG Class: BAC ends.

```
FEATURES
    source      location/Qualifiers
1. 461
    /organism="Oryza rufipogon"
    /mol_type="genomic DNA"
    /db_xref="taxon:4529"
    /clone="OR_CBA0004J23"
```

```

/issue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_Cha"
/notes="vector: pGIRAC1; site_1: HindIII; site_2: HindIII
dirk treated 36 hrs before harvest"

```

[illegible][illegible]







source  
1. .618  
/organism="Oryza rufipogon"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4529"  
/clone\_lib="B84098B05"  
/issue\_type="Young Leaves"  
/lab\_host="DH10B-T1 phage resistant"  
/clone\_lib="OR B84"  
/note="Vector: pAGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

ORIGIN

Query Match 22.8%; Score 303.4; DB 9; Length 618;  
Best Local Similarity 97.8%; Pred. No. 1e-55;  
Matches 318; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1001 TTTGTTCTTTGTGTAAGTG -CCGGCTCATCGTCCCTGATCATCTTTGTTCTTCTACATC 1059  
Db 610 TTTGTTCTTTGTGTAAGTGCCCGGCTCATCGTCCCTGATCATCTTTGTTCTTCTACATC 551

Qy 1060 TCACTAATGTAACATACATAGATCAGATGCCAAGAGCAGGAGTTGGGGATTAAAGGC 1119  
Db 550 TCATTAATGTAACATACATAGATCAGATGCCAAGAGCAGGAGTTGGGGATTAAAGGC 491

Qy 1120 GAATAGTAAGTAATTTTGTGCTGCTTTGCAAGTATCATCAGTACACCCGGTGAAG 1179  
Db 490 GAATAGTAAGTAATTTTGTGCTGCTTTGCAAGTATCATCAGTACACCCGGTGAAG 431

Qy 1180 CTTAGCTTCCAAATGTGATGTAATTAGACAGCGCCTTCCGTAAGTGTGGCGCCGATGA 1239  
Db 430 CTTAGCTTCCAAATGTGATGTAATTAGACAGCGCCTTCCGTAAGTGTGGCGCCGATGA 371

Qy 1240 TGATCTTGACAGGGGTGCAATTAAGGATTTGATTTCCATTTTGTGATGTAATTGGCAA 1299  
Db 370 TGATCTTGACAGGGGTGCAATTAAGGATTTGATTTTCCATTTTGTGATGTAATTGGCTTA 311

Qy 1300 CTGTCTCATTGGACCAAAAAAAAA 1324  
Db 310 CTGTCTCATTGGACCAAAACATATA 286

RESULT 7  
CG257740/c 884 bp DNA linear GSS 25-AUG-2003  
LOCUS OGC048TH ZM 0.7\_1.5\_KB zea mays genomic clone ZMWBM0650G23,  
DEFINITION genomic survey sequence.  
ACCESSION CG257740  
VERSION CG257740.1 GI:34164524  
KEYWORDS GSS.  
SOURCE zea mays  
ORGANISM zea mays  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 884)  
Whiteley, C.A., Quackenbush, J., Van Aken, S., Utecherack, T.,  
Renwick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other GSSs: OGC048TV  
Contact: Cathy Whiteley  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteley@tigr.org  
Seq primer: TR  
Class: sheared ends.  
location/Qualifiers  
1. .884  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"

FEATURES  
source

/db\_xref="taxon:4577"  
/clone="ZMWBM0650G23"  
/clone\_lib="ZM 0.7\_1.5\_KB"  
/note="Vector: pBCK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN

Query Match 20.6%; Score 274.4; DB 9; Length 884;  
Best Local Similarity 71.4%; Pred. No. 2.3e-49;  
Matches 528; Conservative 0; Mismatches 96; Indels 116; Gaps 8;

Qy 290 GGGAGGAGTCTCCGGCGCCGCCGACGAGCCGAGCCGCGCGCTTCAAGGAC 349  
Db 869 GAGGAGTGTCTCCGCGGAGCCCGAGGGCTCCGATCCGCGCTCCGCCCAAGGAC 810

Qy 350 GGCGGA-----CACCGCGCCACAGACCGCAAGGGGGCGCGC 388  
Db 809 GCTGAGGAGGAGGAGGAGGCGGTGGGCACGAGCCGAGAACGAGAGGGGCTCGC 750

Qy 389 GGTACGACGGCGGCAAGAGCTGCTCCAAAGCGGAGCGAGAGAGCGCAAGGATCCCA 448  
Db 749 GTT---GACGTGGGAGAGTAGCTTAAGCTAAGCTAAGACGAGCGGAGGATCCCG 693

Qy 449 TGAAGGAGAGAGCAAGGC----- 470  
Db 692 TGAAGGAGAGAGCAAGGTTCTAGAGATCTCCACGTTCTGCTTCAAAAGCT 633

Qy 471 -----TGTTCTGCTGGGCTGGAAGATTGCGCAAGGGG 506  
Db 632 TACACCTTTTGTGGGTGAGGTGTTCTTGTGGGTGAGGATTTGCGCAAGGGGCG 573

Qy 507 ACTGGCGGAGCATCTCGGCAACTTGGTCAATCTGCGGAGCGCAAGCGAGGTGGCGAGC 566  
Db 572 ACTGGCGGAGCATCTCGGCAACTTGGTCAATCTGCGGAGCGCAAGCGAGGTGGCGAGC 513

Qy 567 ACGCGAGAGTACTTCTTCACTCCGCTCAACTCATGAAACCGGACCGCGCGCTTCAAGCA 626  
Db 512 ACGCGAGAGTACTTCTTCACTCCGCTCAACTCATGAAACCGGAGCGCGCGCTTCAAGCA 453

Qy 627 TCCACGATATCAAGGCTCACCGCGCGGATCAAGTGGCGCGGACGAGGGCGCCCGCA 686  
Db 452 TCCACGATATCAAGGCTGAGCGCGGCGGAGTGGCGCGGCG-----CGGCGCACCA 399

Qy 687 TCACGGCGCACCAAGGCGCAACCCCGGCGGCGGCGCTGGGCGCGCGCGGCGCATGA 746  
Db 398 TCACGGCGCGCCCAACAGG-----CGGAGCGATGCGATGATGATCCCGGCGCATGA 345

Qy 747 AGCACACACACACACACCGCGCGGCGCGCGCGCGCGCATGATGACGCGCG 806  
Db 344 AGCACACACACACACAC-----CGGCTCGCGGATGGGATGATGACGCGCGCG 297

Qy 807 CGCCCATGGGCGCACCGCGTGGCGCGGCAATGTT---GCCCGCGCGTGGCGACGCGCG 863  
Db 296 CGCCCATGGGCGCACCGCGTGGCGCGGCAATGTTGGCGCGCGCGTGGCGACGCGCG 237

Qy 864 TGGTGTCCCGCGGCGG---CCAGCGCGGTACGTCGTCGCGGCTGACCGCGCGCTC 920  
Db 236 TCATGTTCCCGCGGCGGTCATCTCGGCTATGTGTCAGATGAGATACCGCGGCGCGC 177

Qy 921 CGGCAAGATGACCAATGA 940  
Db 176 CGGCAAGATGACCAATGA 157

RESULT 8  
CC710311/c 689 bp DNA linear GSS 19-JUN-2003  
LOCUS OGDH26TC ZM 0.7\_1.5\_KB zea mays genomic clone ZMWBM0145F04,  
DEFINITION genomic survey sequence.  
ACCESSION CC710311  
VERSION CC710311.1 GI:32115087  
KEYWORDS GSS.  
SOURCE zea mays

ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 689)  
REFERENCE Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
AUTHORS Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Numburg,A., Robbins,D. and Lakey,N.  
TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Other GSSs: OGDH26TM  
CONTACT: Cathy Whitelaw  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.

FEATURES  
source  
1. .689  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZM0145F04"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 19.9%; Score 264.6; DB 9; Length 689;  
Best Local Similarity 71.9%; Pred. No. 3.1e-47;  
Matches 429; Conservative 0; Mismatches 134; Indels 34; Gaps 5;

466 CAGGCTTTCTGCTGGGCTGACAAAGTTGCGCAAGGGGACTGGCGAGCATCTGCG 525  
|||  
576 CAGGTTTCTTGTGTGGCTTGGACAAAGTTGCGCAAGGGGACTGGCGAGCATCTGCG 517  
|||  
526 CAACTTGATCTCGCGGACGCGCAAGGTGGCGACGACGCGAGAAATCTTCAT 585  
|||  
516 CAACTTGATCTCGCGGACGCGCAAGGTGGCGACGACGCGAGAAATCTTCAT 457  
|||  
586 CCGGCTTAATCTCAATGAACCGGACCGCGCTCCAGATCCAGACATCACCAGCT 645  
|||  
456 CCGGCTTAATCTCAATGAACCGGACCGCGCTCCAGATCCAGACATCACCAGCT 397  
|||  
646 CACGCGCGGATGAGTCCGCGCGACGAGGCGCCCGATCACCGGACAGGCGAC 705  
|||  
396 GACGCGCGGATGAGTCCGCGCGACGAGGCGCCCGATCACCGGCGGCGCGCGC 337  
|||  
706 GAGGCAACCCGCGCGGCGGCTGGGCGCGCGCGGATGAGGACGACGACGACCA 765  
|||  
336 GGG-----GGCGATGCGGATGGGGCGGGCGGATGAGGACGACGACCA 293  
|||  
766 CCGGCGCGGCGCGCGCGCGCGCGGATGAGGCGCGCGCGGATGAGGCGCGCG 825  
|||  
292 -----CGGTTCCGCGCGATGGGCGATGTAAGGCGCGCGCGGCGCGCGGT 244  
|||  
826 CGCGCGCGCATGGT---GCCCGCGCGGTCGCGGACGCGCGGTGGTTCGCGCGCGCG 880  
|||  
243 CGCGCGCGCATGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 184  
|||  
881 -CAGCGCGCGTACGTCGTCGCGCGGCTACCGCGCGCTCCGCGCAAGATGACCAATG 939  
|||  
183 TCACTCGCGCTAATCTCTGCGCGGTGAGATACCGCGCGCGCGCGCAAGATGACCAATG 124  
|||  
940 ACGCGCATGACGCG-ACATGACGATTTCTTCTCTCTCTTTCTTGATGATCATCTT 998  
|||  
123 ATCTGTACGACCGCGCATTTGGCGAGCTGGCGACGCTCATCTCTTGTCTCTCTAC 64  
|||  
999 GATTGTCTTTGTGTAGTGGCGGCTCATGTCCTCGATCATTTGTCTCTCTAC 1055  
|||

Db 63 GATTGCTGATGATCTTGTGTAGCCAAAGCTATCTATCCATCTTCTCATCTC 7

RESULT 9  
CG366978/c 804 bp DNA linear GSS 26-AUG-2003  
LOCUS OGMAM96TV ZM 0.7\_1.5\_KB Zea mays genomic clone ZMBMA0520023,  
DEFINITION genomic survey sequence.  
ACCESSION CG366978  
VERSION CG366978.1 GI.34284245  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 804)  
REFERENCE Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
AUTHORS Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Numburg,A., Robbins,D. and Lakey,N.  
TITLE Consortium for Maize Genomics  
JOURNAL Unpublished (2002)  
COMMENT Other GSSs: OGMAM96TV  
CONTACT: Cathy Whitelaw  
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.

FEATURES  
source  
1. .804  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZM014520023"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 19.7%; Score 262.2; DB 9; Length 804;  
Best Local Similarity 74.9%; Pred. No. 1.1e-46;  
Matches 408; Conservative 0; Mismatches 103; Indels 34; Gaps 5;

466 CAGGCTTTCTGCTGGGCTGAGCAAGTTGCGCAAGGGGACTGGCGAGCATCTGCG 525  
|||  
689 CAGGTTTCTTGTGTGGCTTGGACAAAGTTGCGCAAGGGGACTGGCGAGCATCTGCG 630  
|||  
526 CAACTTGATCTCGCGGACGCGCAAGGTGGCGACGACGCGAGAAATCTTCAT 585  
|||  
629 CAACTTGATCTCGCGGACGCGCAAGGTGGCGACGACGCGAGAAATCTTCAT 570  
|||  
586 CCGGCTTAATCTCAATGAACCGGACCGCGCTCCAGATCCAGACATCACCAGCT 645  
|||  
569 CCGGCTTAATCTCAATGAACCGGACCGCGCTCCAGATCCAGACATCACCAGCT 510  
|||  
646 CACGCGCGGATGAGTCCGCGCGACGAGGCGCCCGATCACCGGACAGGCGAC 705  
|||  
509 GACGCGCGGATGAGTCCGCGCGACGAGGCGCCCGATCACCGGCGGCGCGCGCGC 450  
|||  
706 GAGGCAACCCGCGCGGCGGCTGGGCGCGCGCGGATGAGGACGACGACGACGACCA 765  
|||  
449 GGG-----GGCGATGCGGATGGGGCGGGCGGATGAGGACGACGACCA 406  
|||  
766 CCGGCGCGGCGCGCGCGCGCGCGGATGAGGCGCGCGCGGATGAGGCGCGCGGT 825  
|||  
405 -----CGGTTCCGCGCGATGGGCGATGTAAGGCGCGCGCGCGCGCGCGGT 357  
|||  
826 CGCGCGCGCATGGT---GCCCGCGCGGTCGCGGACGCGCGGTGGTTCGCGCGCGCG 880  
|||

Db 356 CGCCGGGACATGGTGGCGCCGGCCGCTCGGACGCGGTCATGTTCCGCGCGGCA 297  
Qy 881 -CAGCGCGCTACGTCGTGCGCTGCTACCGCGCGCTCGCGGCAAGATGACCAATG 939  
Db 296 TCATCGCCCTTATGCTCTGCGGTGGATACCGCGCGCACCGACCAAGATGACCAATG 237  
Qy 940 ACGCGCCATGACGG-ACATGAGCAGATTTCTTCTCTCTTCTTCTTATGTCATCTT 998  
Db 236 ATCGTCTACGACGCGCATTTGGCGAGCTGCGCAGCTGCATCCCTTGTCTGTTCTT 177  
Qy 999 GATTT 1003  
Db 176 GATTT 172

RESULT 10  
LOCUS CG362076 853 bp DNA linear GSS 26-AUG-2003  
DEFINITION OG1CL41TH ZM 0.7-1.5\_KB Zea mays genomic clone ZMMBMA0730H10,  
genomic survey sequence.  
ACCESSION CG362076  
VERSION CG362076.1 GI:34279343  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 853)  
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other\_GSSs: OG1CL41TV  
Contact: Cathy WhiteLaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteLaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source location/Qualifiers  
1..853  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0730H10"  
/clone\_lib="ZM 0.7-1.5\_KB"  
/note="Vector: pBCSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 19.7%; Score 262.2; DB 9; Length 853;  
Best Local Similarity 74.9%; Pred. No. 1.1e-46;  
Matches 408; Conservative 0; Mismatches 103; Indels 34; Gaps 5;

Qy 466 CAGGCTGTTCTTGTGGGCTGGAACAAGTTGGGAAGGGGACTGCGGAGCATCTCGG 525  
Db 264 CAGGTTGTTCTTGTGGGCTGGAACAAGTTGGGAAGGGGACTGCGGAGCATCTCGG 323  
Qy 526 CAATTTGTCTCTTCGGGAGCCCAAGCAGAGTGGCGAGCCAGCGCAAGTACTTCTAT 585  
Db 324 CAATTTGTCTCTTCGGGAGCCCAAGCAGAGTGGCGAGCCAGCGCAAGTACTTCTAT 383  
Qy 586 CCGGCTCACTCATGTAACCGGACCGCGGCTCGCATCCAGATCAGCATCAGAGGT 645  
Db 384 CCGGCTCACTCATGTAACCGGACCGCGGCTCGCATCCAGATCAGCATCAGAGGT 443  
Qy 646 CACGCGCGGATCATGTTGCGCGGACAGAGGCGCCCGATCACCGGCCACAGGCCAC 705  
|||||

Db 444 GACCGCGGACAGGTGGCCACGCGCGCGCCCATCAACGGACGCGCGCCGACCGCGC 503  
Qy 706 GGGCAACCCCGCGCGCGCGCTGCGCGCGCGCGGAGTGAAGACACACACCAACCA 765  
Db 504 GGG-----GGGATGCCGATGGGCGCGCGCGGATGAAGACACCAAC----- 547  
Qy 766 CCGGCGGCGCGCGCGCGCCCATGCCATGATGACGCGCGCGCCCATGAGGACACCCCGT 825  
Db 548 -----CGGTTCCGCGGATGGGATGTACGGGACGCGCCCATGGGACACCCGCT 596  
Qy 826 CGCGGCGCATGTTGTT---GCCCGCGCGCTGGGACGCGCGGTTGTTCCGCGCGGCG- 880  
Db 597 CGCGGCGCATGTTGTTGCGCGCGCGCGCGCTGCGCACCGCGGTCAATGTTCCCGCGGCGCA 656  
Qy 881 -CAGCGCGCTTACGTCGTCGCGCTACCGCGCGCTCGCGGCAAGATGACCAATG 939  
Db 657 TCATCGCCCTTATGCTCTGCGGTGGATACCGCGCGCACGACCAAGATGACCAATG 716  
Qy 940 ACGCGCCATGACGG-ACATGAGCAGATTTCTTCTCTCTTCTTCTTATGTCATCTT 998  
Db 717 ATCGTCTACGACGCGCATTTGGCGAGCTGCGCAGCTGCATCCCTTGTCTGTTCTT 776  
Qy 999 GATTT 1003  
Db 777 GATTT 781

RESULT 11  
LOCUS CG219911/c 933 bp DNA linear GSS 22-AUG-2003  
DEFINITION OG1AD74TH ZM 0.7-1.5\_KB Zea mays genomic clone ZMMBMA0715N04,  
genomic survey sequence.  
ACCESSION CG219911  
VERSION CG219911.1 GI:34119799  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 933)  
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other\_GSSs: OG1AD74TV  
Contact: Cathy WhiteLaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whiteLaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source location/Qualifiers  
1..933  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0715N04"  
/clone\_lib="ZM 0.7-1.5\_KB"  
/note="Vector: pBCSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 19.7%; Score 262.2; DB 9; Length 933;  
Best Local Similarity 74.9%; Pred. No. 1.1e-46;  
Matches 408; Conservative 0; Mismatches 103; Indels 34; Gaps 5;

Qy 466 CAGGCTGTTCTTGTGGGCTGGAACAAGTTGGGAAGGGGACTGCGGAGCATCTCGG 525  
|||||

Db 786 CAGGTTCTTGTGGTGTGACGAAGAAGTTCCGGAGAGGCGACTGGCGGACATCTCGCG 727

Qy 526 CAACCTTGTCATCTCGCGAGACGCCAAGCAGAGTGGCGAGCCACGCGAGAAGTACTTCAT 585

Db 726 CAACCTTGTCATCTCGCGAGACGCCAAGCAGAGTGGCGAGCCACGCGAGAAGTACTTCAT 667

Qy 586 CCGGCTCACTCATGACGAGACCGCGAGCCCGCGCTCCAGAGATCCAGACATCCACGCGT 645

Db 666 CCGGCTCACTCATGACGAGACCGCGAGCCCGCGCTCCAGAGATCCAGACATCCACGCGT 607

Qy 646 CACCGCGCGAGATCAGGTGCGCGCGAGAGGCGCGCCGATCACCAGCCACAGCGCAC 705

Db 606 GACCGCGCGAGAGTGGCGAGCGCGCGCGCGCCCATACAGGCGCGCGCGCGCGCGCG 547

Qy 706 GGGCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTAAAGACACCAACCA 765

Db 546 GGG-----GGCGATGCGGATGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 503

Qy 766 CCGGCG 825

Db 502 -----CGGTTCCGCGCGAGTGGCGAGTGAAGGCGCGCGCGCGCGCGCGCGCGCG 454

Qy 826 CCGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 880

Db 453 CCGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 394

Qy 881 -CAGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 939

Db 393 TCACGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 334

Qy 940 ACGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 998

Db 333 ATGCTTACACGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 274

Qy 999 GATT 1003

Db 273 GATT 269

RESULT 12

CL968676 807 bp DNA linear GSS 21-SEP-2004

LOCUS OeIFCC017023 Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.

ACCESSION CL968676

VERSION CL968676.1 GI:52391986

KEYWORDS GSS.

SOURCE Oryza sativa (indica cultivar-group)

ORGANISM Oryza sativa (indica cultivar-group)

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

AUTHORS 1 (bases 1 to 807)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.

TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis

JOURNAL Unpublished (2004)

COMMENT Contact: Chen Chen

Department of Bioinformatics

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-8048676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

FEATURES

source

1. .807

Location/Qualifiers

/organism="Oryza sativa (indica cultivar-group)"

/mol\_type="genomic DNA"

/db\_xref="taxon:39946"

ORIGIN

Query Match 19.6%; Score 260.6; DB 9; Length 807;

Best Local Similarity 83.9%; Pred. No. 2, 4e-46;

Matches 323; Conservative 0; Mismatches 49; Indels 13; Gaps 2;

Qy 362 CGGACGAGCGCAGAGCG 421

Db 436 CGTGGCGAAGAGCGGATGTCGTGCGAGGTACGTGCGCAAGATCTGCTCCAGTCG 495

Qy 422 GAGCAGAGAGCG 481

Db 496 GAGCAGAGAGCG 555

Qy 482 GGGCTGACAAAGTTCGCGCAAGGAGGAGTGGCGAGCATCTCGCGCAACTTCGATCTCG 541

Db 556 GGGCTGACAAAGTTCGCGCAAGGAGGAGTGGCGAGCATCTCGCGCAACTTCGATCTCG 615

Qy 542 CGGACGCGCAAGCGAGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 601

Db 616 CGGACG 675

Qy 602 AACCG 661

Db 676 AACCG 732

Qy 662 GTCCGCGCGCGAGAGGCG 721

Db 733 GTTCCGCGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 782

Qy 722 GCGCGCGTGGCG 746

Db 783 GCGCGCGTGGCG 807

RESULT 13

CC751932

LOCUS ZMMBB0133C24.1 ZMMBB Zea mays genomic clone ZMMBB0133C24.3, genomic survey sequence.

ACCESSION CC751932

VERSION CC751932.1 GI:32211515

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 740)

Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bhatti, A.K., Messing, J. and Wing, R.

TITLE Sequencing of the maize genome

JOURNAL Unpublished (2003)

COMMENT Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: T7

BACKWARD: M13r

Plate: 0133 row: C column: 24

Seq primer: M13r

Class: BAC ends.

FEATURES

source

1. .740

Location/Qualifiers

/organism="Zea mays"

/mol\_type="genomic DNA"



Fax: 520 621 9288  
Email: <http://genome.arizona.edu>  
PCR Primers  
FORWARD: T7  
BACKWARD: M13r  
Plate: 0133 row: P column: 10  
Seq primer: T7  
Class: BAC ends.

FEATURES  
Source

Location/Qualifiers  
1..653  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBB01.3P10"  
/lab\_host="DH108"  
/clone\_lib="ZMMBBb"  
/note="Vector: pBelBAC1; Site\_1: HindIII; Site\_2:  
HindIII; Zea mays L. ssp. mays"

## ORIGIN

Query Match 18.9%; Score 251.2; DB 9; Length 653;  
Best Local Similarity 73.8%; Pred. No. 2,7e-44;  
Matches 400; Conservative 0; Mismatches 108; Indels 34; Gaps 5;

```
QY 469 GCTGTTCTTGTGCTGGGCTGACAAAGTTCCGCAAGGGGACTGGCGAGCATTTCCGCAA 528
Db 63 GTTGTCTTGTGTGGTCTGGACAAAGTTCCGCAAGGGGACTGGCGAGCATTTCCGCAA 122
QY 529 CTTTCGTATCTCGGGAGCGCAAGCGAGTGGCGAGCCAGCGGAGAAATCTTCATCCG 588
Db 123 CTTTCGTATCTCGGGAGCGCGAGCGAGTGGCGAGCCAGCGGAGAAATCTTCATCCG 182
QY 589 CTTCAACTCATGAACCGGACCGCGCGCTCCAGCATCCAGCATCAACAGCTCAC 648
Db 183 CTTCAACTCATGAACCGGACCGGCGCGCTCCAGCATCCAGCATCAACAGCTCAC 242
QY 649 CGCGGCGATCAGTCCCGCGCAGCAGGGCGCCCATCAACGGCCACCAAGCCACGGG 708
Db 243 CGCGGCGAGTGGCCACCGCGCGCGCCCATCAACGGCGCGCGCCGACCGCGCGGG 302
QY 709 CAACCCCGCGCGCGCGCTGGCGCGCGCGCGCATGAAGACCAACCAACCAACCC 768
Db 303 -----GCGATGCCGATGGGGCGGGCGCGCATGAAGACCAACCAACCC----- 343
QY 769 GGGCGGCGCGCGCGCGCATGCCATGTAACAGCGCGCGCGCATGGGCCACCCGTCGC 828
Db 344 -----CGGTTCCGCGCATGGGCATGTACGGGCAACCGGCCATGGGCCACCGGTCGC 395
QY 829 CGGCCACATGGT--GCCCGCGCGCGTCCGACCGCGGTGTTCCTCCGCGGGC---CA 882
Db 396 CGGGACATGGTGGCGCGCGGGCGCGTCCGACCGCGGTATGTTCCCGCGGGCCATCA 455
QY 883 CGCGCGGTACGTCTGCGCGCGCTACCGCGCGCTCCCGGCCAAGATGACCAATAGC 942
Db 456 CTGCGCTATGTCTGCGCGGTATACCGGGCGCACCAACGACGACGCGATC 515
QY 943 CGCCATGACGG--ACATGACGAGCATTTCTCTCTCTTCTTGAATGTCATCTTGAT 1001
Db 516 GCCTACGACCGGCAATTCCGCGAGCTGGCGCACTGCATCCCTTGTGCTGTTCTTTGAT 575
QY 1002 TT 1003
Db 576 TT 577
```

Search completed: July 29, 2005, 22:29:27  
Job time : 4781 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 15:06:02 ; Search time 785 seconds  
(without alignments)  
10029.625 Million cell updates/sec

Title: US-10-630-636-1

Perfect score: 1330  
Sequence: 1 gtgcgagatccaccaccga.....gacccaaaaaaaaaaaaaa 1330

Scoring table: IDENTITY NUC  
Gap0 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1330	100.0	1330	12	AD000777 Rice Myb
2	918	69.0	918	12	AD000780 Rice Myb
3	135.4	10.2	1161	3	AAC34213 Arabidops
4	130.6	9.8	412	12	ADP93595 Cotton ex
5	129.2	9.7	1432	10	ADCA6636 Thalecres
6	129.2	9.7	1432	10	ADDS5829 Thalecres
7	129.2	9.7	1432	12	AD161458 CDNA enco
8	129.2	9.7	1432	12	AD002450 Thalecres
9	126.8	9.5	1119	6	ABT08115 Myb-relat
10	126.8	9.5	1119	12	ADH50105 Soybean M
11	123.2	9.3	257	12	ADQ03867 Maize tra
12	120.2	9.0	386	12	ADQ03936 Maize tra
13	119.8	9.0	316	12	ADQ03772 Maize tra
14	119	8.9	144	12	ADQ03861 Maize tra
15	118.6	8.9	867	6	ABK65275 Arabidops
16	118.6	8.9	867	10	AD030621 Plant yie
17	118.6	8.9	867	10	AD031552 Plant yie
18	118.6	8.9	867	12	AD144216 Plant tra
19	112.8	8.5	1374	12	AD143214 Plant tra
20	112.8	8.5	1428	12	AD143213 Plant tra

21	111.8	8.4	934	3	AAC57273 Eucalyptu
22	111.8	8.4	1358	3	AAC56101 Eucalyptu
23	110.6	8.3	954	12	AD000782 Rice Myb
24	110.6	8.3	1287	12	AD000779 Rice Myb
25	109.2	8.2	1009	10	ADP38031 Synchroni
26	108	8.1	866	6	ABK65370 Arabidops
27	108	8.1	866	10	ADDS0950 Plant yie
28	108	8.1	866	12	ADB37148 Plant yie
29	108	8.1	866	12	AD141988 Plant tra
30	108	8.1	866	12	AD003516 Thalecres
31	108	8.1	1169	3	AAC47819 Arabidops
32	107.6	8.1	804	10	ADD29922 Plant yie
33	107.6	8.1	804	12	AD144406 Plant tra
34	107	8.0	1277	13	ADRE0712 Cotton CD
35	106.4	8.0	349	3	AAC56698 Eucalyptu
36	106.4	8.0	828	12	AD000781 Rice Myb
37	106.4	8.0	1012	12	AD000778 Rice Myb
38	105.6	7.9	335	3	AAC56762 Eucalyptu
39	105	7.9	422	3	AAC56944 Pinus rad
40	105	7.9	436	3	AAC57318 Pinus rad
41	104.8	7.9	632	3	AAC56528 Eucalyptu
42	104.4	7.8	274	12	AD006568 Soybean t
43	103.8	7.8	589	13	ACN46091 Cotton pr
44	103.8	7.8	598	13	ACN54935 Cotton an
45	102.2	7.7	284	12	ADQ06567 Soybean t

## ALIGNMENTS

RESULT 1  
AD000777  
ID AD000777 standard; cDNA; 1330 BP.  
XX  
AC AD000777;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Rice Myb protein OsmYBS1 cDNA.  
XX  
OS  
XX  
KW ss; gene; OsmYBS1; rice; plant; transgenic; Myb; transcription factor.  
XX  
OS Oryza sativa.  
XX  
FH Key Location/Qualifiers  
FT CDS 20..940  
FT /tag= a  
FT /product= "OsmYBS1"  
XX  
XX US2004107456-A1.  
XX  
XX 03-JUN-2004.  
XX  
XX 30-JUL-2003; 2003US-00630636.  
XX  
XX 31-JUL-2002; 2002US-0399999P.  
XX  
XX (YUUS//) YU S.  
XX  
XX YU S;  
XX  
XX WPI: 2004-419531/39.  
XX  
XX P-PSDB: AD000783.  
XX  
XX Novel Myb proteins derived from Oryza sativa and designated OsmYBS1,  
XX OsmYBS2 or OsmYBS3; useful for regulating expression of a gene in plant  
XX cell, preferably rice or barley cell.  
XX  
XX Claim 18; SEQ ID NO 1; 22pp; English.  
XX  
XX The invention relates to a pure polypeptide comprising a fully defined  
XX Myb protein transcription factor OsmYBS1, OsmYBS2 or OsmYBS3 protein  
XX sequence, where the polypeptide regulates expression of a gene in a cell.

The polypeptide is useful for regulating expression of a gene in a plant cell, preferably rice or barley cell. The method is useful for identifying a compound that modulates the activity of the polypeptide. An antibody to the polypeptide is useful for detecting the presence and distribution of OsMYBS proteins in tissues and in cellular compartments and for verifying the expression of OsMYBS proteins in a transgenic plant. Nucleic acid encoding the polypeptide is useful for generating a transgenic plant or producing to determine whether an OsMYBS mRNA is expressed in a tissue or cell. Nucleic acid encoding the polypeptide is also useful as primers in PCR-based detection methods, or as labelled probes in nucleic acid blots (e.g. Northern blots). The compound identified by the method is useful for enhancing or repressing the expression of a gene regulated by the polypeptide, e.g. alpha-amylase gene. The present sequence represents cDNA encoding the rice *Myb* protein OsMYBS1.

**SQ** Sequence 1330 BP; 254 A; 438 C; 435 G; 203 T; 0 U; 0 Other;

Query Match	Score	DB	Length
100.0%	1330	12	1330

Matches 1330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTGGGAGATTCACCAACCCGATGACCTCCAGGCGGCGACGAGACGACCAACGCGCGCGCG	60
Dp	1	GTGCGAGATTCACCAACCCGATGACCTCCAGGCGGCGACGAGACGACCAACGCGCGCGCG	60
QY	61	GCGCGCGCGCGTGGACCAAGGAGAGACGACCAAGGCGTTTCGAAACGCGCTTCGCGCTTCGCG	120
Dp	61	GCGCGCGCGCGTGGACCAAGGAGAGAGACCAAGGCGTTTCGAAACGCGCTTCGCGCTTCGCG	120
QY	121	GCGCGCGCGCGCGCGCGGAGACGAGGCGCGCGCCGACGACGACTGGTTTGCCTGCGCTCGCGC	180
Dp	121	GCGCGCGCGCGCGCGCGGAGACGAGGCGCGCGCCGACGACGACTGGTTTGCCTGCGCGCTCGCGC	180
QY	181	GAGCGTGCCTCCGGGAGGAGGTTCGCGCGGAGAGGTTCGCGGAGGACCTACGAGGCGCTGGTGA	240
Dp	181	GAGCGTGCCTCCGGGAGGAGGTTCGCGCGGAGAGGTTCGCGGAGGACCTACGAGGCGCTGGTGA	240
QY	241	GGACGTCGCGGCGCATTCGACGCGCGGCGCGCTCCGCTCCGCGCTACGCGCGGAGGAGTTC	300
Dp	241	GGAAGTCGCGGCGCATTCGACGCGCGGCGCGCTCCGCTCCGCGCTACGCGCGGAGGAGTTC	300
QY	301	CGCGGCGCGCGCGCGACGAGACCTCGAGACCGCGCGCGCGCTCCAMGAGTCGGGAGACCG	360
Dp	301	CGCGGCGCGCGCGCGACGAGACCTCGAGACCGCGCGCGCGCTCCAMGAGTCGGGAGACCG	360
QY	361	GCGCGGACGAGGCGCAGGGCGCGCGGCGCGGCTTACGACGCGCGGCGCAAGAGTGTCCAAAGG	420
Dp	361	GCGCGGACGAGGCGCAGGGCGCGCGGCGCGGCTTACGACGCGCGGCGCAAGAGTGTCCAAAGG	420
QY	421	GGAGCAGAGAGGCGCAAGGGCGCATCCCATGGACGAGAGAGCAGCGCTGTTCTTGGCT	480
Dp	421	GGAGCAGAGAGGCGCCCAAGGGCGCATCCCATGGACGAGAGAGCAGCGCTGTTCTTGGCT	480
QY	481	GGGGCTGGACAAATTTCGGCGAAGGGGGACTGGCGGAGCATTCGCGGCACTTCGCTCATTC	540
Dp	481	GGGGCTGGACAAATTTCGGCGAAGGGGGACTGGCGGAGCATTCGCGGCACTTCGCTCATTC	540
QY	541	GCGGACGCGCAACGCAAGGTGGCGAGCGACGCGCGAGAAATCTTCATCCGCTCAACCTCAAT	600
Dp	541	GCGGACGCGCAAGCAAGGTGGCGAGCGACGCGCGAGAAATCTTCATCCGCTCAACCTCAAT	600
QY	601	GAACCGCGACCGCGCGCTTCAGCATCCACGACATCAACAGGTCACCGCGCGCGCATTA	660
Dp	601	GAACCGCGACCGCGCGCTTCAGCATCCACGACATCAACAGGTCACCGCGCGCGCATTA	660
QY	661	GATCGCGCGCGACGAGGGCGCGCTCCGATCACTCGGCGACGAGGCCACGCGGCAACCCCGCGG	720
Dp	661	GATCGCGCGCGACGAGGGCGCGCTCCGATCACTCGGCGACGAGGCCACGCGGCAACCCCGCGG	720
QY	721	GCGCGCGCTGGGCGCGCGGCGCATGAAGACCAACCAACCAACCAACCCCGGCGCGCGCGC	780
Dp	721	GCGCGCGCTGGGCGCGCGGCGCATGAAGACCAACCAACCAACCAACCCCGGCGCGCGCGC	780

QY	781	GGCGCCCATGCCCATGATGACAGCGCCGGGCCCATGAGGCAACCCGTCGCGCGGCACATGAT	840
Db	781	GGCGCCCATATCCCATGTACACGCGCCGCCCATAGGGCACCCCGTCGCGCGGCACATGAT	840
QY	841	GGCGCGCGCCGTCGAGCACGCGCGGTGTGTTCGCCCGGGGACACGCGCGTCATGTCGTGCC	900
Db	841	GGCGCGCGCGGT	900
QY	901	CGTGGGCTATCCCGGCGCTTCGCGCAAGATGACCAATGACGGGCCATGAGCGGACATGA	960
Db	901	CGTCGGCTATCCCGGCGCTTCGCGCAAGATGACCAATGACGGGCCATGAGCGGACATGA	960
QY	961	GGAGCATTTCTTCCTCCTCTTCTTGTGATGTCAATCTGATTTGTTGTTGTGTGATGCGC	1020
Db	961	GGAGCATTTCTTCCTCCTCTTCTTGTGATGTCAATCTGATTTGTTGTTGTGTGATGCGC	1020
QY	1021	CGGCTCATCGTCCCTGATCATCTTGTGTCCTGACATCTCAGTAATGTAACATACATATA	1080
Db	1021	CGGCTCATCGTCCCTGATCATCTTGTGTCCTGACATCTCAGTAATGTAACATACATATA	1080
QY	1081	GATCAGATGCCAAGATGACAGGATTTGGGGATTTAAGGCGAATPAAGTAAGTATTTTGGCT	1140
Db	1081	GATCAGATGCCAAGATGACAGGATTTGGGGATTTAAGGCGAATPAAGTAAGTATTTTGGCT	1140
QY	1141	GATCGTTTGAAGGATCATACAGTACACCCGGGAAAGCTTATGCTCCAAATGATGATGT	1200
Db	1141	GATCGTTTGAAGGATCATACAGTACACCCGGGAAAGCTTATGCTCCAAATGATGATGT	1200
QY	1201	AATTAGCAGCGGCGCTTCCTGACGTGTGGCGCGATGATGATCTTGGCAGGGTTGGCAAT	1260
Db	1201	AATTAGCAGCGGCGCTTCCTGACGTGTGGCGCGATGATGATCTTGGCAGGGTTGGCAAT	1260
QY	1261	TAGGGATTTGATTTTCATTTTGTCTGATGTAAATTTGCCAATGTCTCATTTGGACCAAAAAA	1320
Db	1261	TAGGGATTTGATTTTCATTTTGTCTGATGTAAATTTGGCAATGTCTCATTTGGACCAAAAAA	1320
QY	1321	AAAAAAAAAA 1330	
Db	1321	AAAAAAAAAA 1330	

XX	RESULT 2
AD000780	ID AD000780 standard; cDNA; 918 BP.
AC	AD000780;
DT	26-AUG-2004 (first entry)
DE	Rice Myb protein OsMYB1 CDS cDNA.
OS	Oryza sativa.
FT	Key Location/Qualifiers
FT	CDS 1..918
FT	/tag= a
FT	/partial
FT	/product= "OsMYB1"
FT	/note= "No stop codon given"
PN	US2004107456-A1.
PD	03-JUN-2004.
PF	30-JUL-2003; 2003US-00630636.
PR	31-JUL-2002; 2002US-039999P.
PA	(YUUS/) YU S.
XX	





PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0136021P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139453P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139452P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151065P.
PR 21-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 24-JUN-1999;	99US-0140695P.	PR 01-SEP-1999;	99US-0151930P.
PR 28-JUN-1999;	99US-0140823P.	PR 07-SEP-1999;	99US-0152363P.
PR 29-JUN-1999;	99US-0140991P.	PR 10-SEP-1999;	99US-0153070P.
PR 30-JUN-1999;	99US-0141287P.	PR 13-SEP-1999;	99US-0153758P.
PR 01-JUL-1999;	99US-0141842P.	PR 15-SEP-1999;	99US-0154018P.
PR 02-JUL-1999;	99US-0142055P.	PR 16-SEP-1999;	99US-0154039P.
PR 06-JUL-1999;	99US-0142330P.	PR 20-SEP-1999;	99US-0154779P.
PR 06-JUL-1999;	99US-0142803P.	PR 22-SEP-1999;	99US-0155139P.
PR 09-JUL-1999;	99US-0142920P.	PR 23-SEP-1999;	99US-0155486P.
PR 12-JUL-1999;	99US-0143542P.	PR 24-SEP-1999;	99US-0155659P.
PR 14-JUL-1999;	99US-0143624P.	PR 28-SEP-1999;	99US-0156458P.
PR 15-JUL-1999;	99US-0144005P.	PR 05-OCT-1999;	99US-0157753P.
PR 16-JUL-1999;	99US-0144085P.	PR 06-OCT-1999;	99US-0157665P.
PR 16-JUL-1999;	99US-0144086P.	PR 07-OCT-1999;	99US-0158029P.
PR 19-JUL-1999;	99US-0144331P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159637P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159638P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159931P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159931P.
PR 21-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159932P.
PR 21-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159932P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 23-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160814P.
PR 23-JUL-1999;	99US-0145192P.	PR 22-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.  
PR 23-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match	10.2%	Score 135.4;	DB 3;	Length 1161;
Best Local Similarity	70.4%;	Pred. No. 8.1e-13;		
Matches 181; Conservative	0;	Mismatches 76;	Indels 0;	Gaps 0

Oy	393	ACGACGGGGCGCAAGACCTGCTCCAGACGGGACAGAGAGGGCGCAAGGCGCATCCCATGGA	452
Db	419	ATGAGAGAAAGATGGGTGGCTGAGAGCTGAGACAGAGAGAAAGAAAGGAGATTCATGGA	478
Oy	453	CGAGGAAGAGCACAAGGCTGTTCTTGCTGGGGCTGGACAAGTTTCGGCAAGGGGGA	512
Db	479	CTGAAGAAAGAGCATCGGTGTGTTCTTTTGGGTTTGGAACAAGTTTGGAAAGAGATTGGA	538
Oy	513	GGAGCATCTCGGCGCACTTCGTCAATCTCGCGGACCGCAACGCAAGTGGCGGAGCCACGGCC	572
Db	539	GAAGCATTTCAAGGAACCTTTGTAATCTCAAGAACTTCAACACAAGTTGGAATGATGCTC	598
Oy	573	AGAAGTACTTCATTCGGCTCAACTCATGAACCGGAGACGGCGCGCTCCAGATTCACAG	632
Db	599	AAAAATACTTCATTCAGGCTTAACGTGATGAAACCGAGATGGAAGGCGGCTTACATTCACAG	658
Oy	633	ACATCACCAAGCGCTACC	649
Db	659	ACATCACCACTGTGAAC	675

## RESULT 4

ADP93595  
ID ADP93595 standard; cDNA; 412 BP.

XX  
AC ADP93595;

XX	
DT	09-SEP-2004 (first entry)

XX Cotton expressed sequence tag, EST, #2606.

XX Cotton; SS, EST; expressed sequence tag; plant; plant protection;; KW

**KW** plant improvement; marker-assisted breeding.  
**XX**

OS Gossypium hirsutum; variety Nucotton33B.  
XX

PN US2004123338-A1.  
XX

PD 24-JUN-2004.  
XX

PF 08-DEC-2000; 2000US-00732627  
XX

PR 10-DEC-1999; 99US-0170255P  
XX

PA (FINC/) FINCHER K L.  
XX

PI Fincher KL;  
XX

DR WPI; 2004-479807/45.  
XX

PT protein or its fragment, useful as molecular tool for the targeting and isolation of novel genes for plant protection and improvement.

XX  
PS Claim 1; SEQ ID NO 2606; 30pp; English.

XX The invention relates to a substantially purified nucleic acid molecule

CC that encodes a cotton protein or its fragment comprising an EST  
CC (expressed sequence tag) appearing as AB090390-ADP95191. Also included  
CC are a substantially purified cotton protein or its fragment encoded by a  
CC nucleic acid molecule above and a transformed plant having a nucleic  
CC acid molecule which comprises: an exogenous promoter region which  
CC functions in a plant cell to cause the production of a mRNA molecule; a  
CC structural nucleic acid molecule comprising one of the ESTs or their  
CC complements; a 3' non-translated sequence that functions in the plant  
CC cell to cause termination of transcription and addition of polyadenylated  
CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as  
CC molecular tool for the targeting and isolation of novel genes for plant  
CC protection and improvement. The ESTs are useful for developing new  
CC strategies for understanding critical plant developmental and metabolic  
CC pathways, for isolating genes and promoters, for identifying and mapping  
CC the genes involved in developmental and metabolic pathways, and for  
CC determining gene function. The cotton nucleic acid molecules are useful  
CC as molecular tags to isolate genetic regions, isolate genes, map genes,  
CC and determine gene function. The nucleic acid molecules are useful for  
CC determining if genes are members of a particular gene family and for use  
CC in marker-assisted breeding programs. The present sequence is one of the  
CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed  
CC in the specification and are available in electronic format from the  
CC USPTO at [seqdata.uspto.gov/sequence.html?IDocID=20040123338](http://seqdata.uspto.gov/sequence.html?IDocID=20040123338).

Sequence 412 BP; 120 A; 83 C; 106 G; 102 T; 0 U; 1 Other;

Query Match	9.8%	Score 130.6;	DB 12;	Length 412;
Best Local Similarity	69.3%	Pred. No. 4.9e-12;		
Matches 192; Conservative	0;	Mismatches 84;	Indels 1;	Gaps 1

QY	336	ACGGCGGCAAGAGCTGCTCCAAAGCGGAGCAGAGAGCGCGCAAGGGCATCCCATAGACGG	455
Db	125	ATGAGGGGAAACAAGATTCGAGGTCGACCAGAAAGAAAGAAAGAAATCCATGSACTG	184
QY	456	AGGAAGACACAGCGCTGTTCTTGTGTGGGGCTGACAAATTCGGCAAGGGGGGACTGGCGGA	515
Db	185	AAGAAAGACATAGGTATTCTTACTGTGTTTAGACAACTTTGGGAAAGAGATTGGAGAA	244
QY	516	GCATCTGGCGCACTTGCTCATCTTCGGGACGCCCAAGCAGGTGGCGAGGCCACGGCGAGA	575
Db	245	GCATTTTCAAGAACTTTGTGATATCGAATCTCCGACCGAGTGGCTGAGCCATGCAAAA	304
QY	576	AGTACTTCATCCGCTCAACTCCATGAAACCGGACCGCGCGCTGCCAGATCCACGAC	634
Db	305	AGTATTTTATACGCTTAATTGATGAATTAGAACCGGCGCGGCTCAGATCCACGACN	364
QY	635	ATCACACAGCGTCACCGCGCGCGATCAGGTGCGCGCGCG	671
Db	365	ATCACAGAGTTAACATAGGGGGATGCTGCTCATCAGC	401

## RESULT 5

ADCC46636  
ID ADCC46636 standard; DNA; 1432 BP.

XX AC ADC46636;

XX 18-DEC-2003 (first entry)  
DT

XX Thalecress transcription factor-like DNA G1362.  
DE

XX Thalecress; transcription factor-like protein; ds; seed trait;  
KW

KW transgenic; gene; plant size; stress tolerance; yield,  
disease resistance; plant.

XX  
OS Arabidopsis thaliana.

XX  
PN US2003093837-A1.

XX 15-MAY-2003.  
PD

XX 01-NOV-2002; 2002US-00286264.  
PF

PR 23-MAR-1999; 99US-0125814P.  
PR 22-MAR-2000; 2000US-00533030.  
XX  
PA (KEDD/) KEDDIE J.  
PA (RIEC/) RIECHMANN J L.  
PA (RATC/) RATCLIFFE O.  
PA (ZHAN/) ZHANG J.  
PA (JIAN/) JIANG C.  
PA (PINE/) PINEDA O.  
PA (HEAR/) HEARD J.  
PA (YUGG/) YU G.  
PA (ADAM/) ADAM L.  
PA (BROU/) BROUN P.  
PA (REUB/) REUBER L.  
PA (PILG/) PILGRIM M.  
PA (SAMA/) SAMAHA R.  
XX  
PI Keddle J, Riechmann JI., Ratcliffe O, Zhang J, Jiang C, Pineda O;  
PI Heard J, Yu G, Adam L, Brown P, Reuber L, Pilgrim M, Samaha R;  
XX  
XX WPI; 2003-765498/72.  
DR P-PSDB; ADC46637.  
XX  
XX  
PT Novel transgenic plant having recombinant polynucleotide encoding  
PT polypeptide that alters trait of transgenic plant when compared with same  
PT trait of another plant lacking recombinant polynucleotide.  
XX  
XX Disclosure; SEQ ID NO 35; 165np; English.  
XX  
XX The invention relates to a transgenic plant having recombinant  
XX polynucleotide (I1) encoding polypeptide comprising at least 6  
XX consecutive amino acids of a sequence chosen from the protein sequence  
XX appearing as ADC46603 - ADC46749 (every second sequence), where  
XX recombinant polynucleotide alters a trait of the seed transgenic plant  
XX when compared with same trait of another plant lacking recombinant  
XX polynucleotide. The proteins are transcription factor-like proteins. Also  
XX included are altering (M1) a trait associated with seed (comprising:  
XX transforming a plant with (I1); selecting the transformed plants; and  
XX identifying a transformed plant with seed having altered trait); altering  
XX (M2) the expression levels of at least one gene of a plant (involving  
XX transforming the plant with (I1) and selecting the transformed plant),  
XX altering (M3) a trait associated with a plant's seed (comprising:  
XX transforming the plant with a recombinant polynucleotide comprising a  
XX nucleotide sequence comprising least 18 consecutive nucleotides of a  
XX sequence appearing as ADC46750 - ADC46766 and selecting the transformed  
XX plant) altering (M4) a plant's trait (involving providing a database  
XX sequence, comparing the database sequence with a polypeptide or a  
XX polynucleotide chosen as detailed above, selecting a database sequence  
XX that needs selected sequence criteria and transforming a database  
XX sequence in the plant) and altering a plant's trait (involving providing  
XX a test polynucleotide, hybridising the test polynucleotide with a  
XX polynucleotide detailed above at low stringency and transforming the  
XX hybridising test polynucleotide in a plant to alter a trait of the  
XX plant). The method (M1) is useful for altering a trait associated with  
XX seed. The method (M2) is useful for altering the expressing levels of at  
XX least one gene of a plant. The method (M3) is useful for altering a trait  
XX associated with a plant's seed. The method (M4) is useful for altering a  
XX plant's trait. The method (M4) is useful for altering a plant's trait  
XX such as seed or plant size, stress tolerance, yield or disease  
XX resistance. The present sequence encodes a transcription factor-like  
XX protein/seed trait altering protein of the invention.  
XX  
XX Sequence 1432 BP; 426 A; 269 C; 351 G; 386 T; 0 U; 0 Other;  
XX  
XX Query Match 9.7%; Score 129.2; DB 10; Length 1432;  
XX Best Local Similarity 70.3%; Pred. No. 8e-12;  
XX Matches 173; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Db 640 TTCTCTCTGTTGGATTAAGTACGGAAAGGTGATGGCTAGCATTTTCCCACTT 699  
Qy 533 GTCATCTCGCGACGCCAAGCAGAGTGGCGAGCCGCGAAGTATTCCGCTC 592  
Db 700 GTAGTAAACAAGAACACCGACCCCAAGTTGGCAGCAGTCAAAAGTATTTTCATTCGCTA 759  
Qy 593 AACTCCATGAACCGCAGCGCCGCCCTCCAGATCCAGCATCACACCGTCAACGCC 652  
Db 760 AATTCAATGAACAAGACAGAGCGCATCAAGCATTCACGACATCACTAGTGTGGCAAC 819  
Qy 653 GCGCAT 658  
Db 820 GCAGAT 825  
RESULT 6  
ADD55829  
ID ADD55829 standard; cDNA; 1432 BP.  
XX  
XX ADD55829;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Thalecress environmental stress-related cDNA #94.  
XX  
XX Thalecress; environmental stress; ss; gene; plant; viral infection;  
XX fungal infection; microbial infection; herbicide resistance; heat; cold;  
XX heavy metal; low light; drought; osmotic stress; salt concentration;  
XX transgenic.  
XX  
XX Arabidopsis thaliana.  
XX  
XX US2003131386-A1.  
XX  
XX 10-JUL-2003.  
XX  
XX 22-OCT-2002; 2002US-00278536.  
XX  
XX 23-MAR-1999; 99US-0125814P.  
XX 22-MAR-2000; 2000US-00532591.  
XX  
XX (SAMA/) SAMAHA R.  
XX (HEAR/) HEARD J.  
XX (JIAN/) JIANG C.  
XX (PINE/) PINEDA O.  
XX (REUB/) REUBER L.  
XX (RIEC/) RIECHMANN J L.  
XX (YUGG/) YU G.  
XX (KEDD/) KEDDIE J.  
XX (RATC/) RATCLIFFE O.  
XX (PILG/) PILGRIM M.  
XX (ADAM/) ADAM L.  
XX (BROU/) BROUN P.  
XX (ZHAN/) ZHANG J.  
XX  
XX Samaha R, Heard J, Jiang C, Pineda O, Reuber L, Riechmann JI.;  
XX Yu G, Keddle J, Ratcliffe O, Pilgrim M, Adam L, Brown P, Zhang J;  
XX WPI; 2003-829608/77.  
XX P-PSDB; ADD55830.  
XX  
XX New recombinant polynucleotide for altering the regulation of gene  
XX expression of plants to modify the plant's traits, particularly the  
XX plant's environmental stress tolerance.  
XX  
XX Claim 14; SEQ ID NO 187; 219pp; English.  
XX  
XX The invention relates to a recombinant polynucleotide that alters a  
XX plant's environmental stress tolerance when compared with the same trait  
XX of another plant lacking the recombinant polynucleotide. Also included  
XX are a transgenic plant comprising the novel recombinant polynucleotide  
XX having a sequence that encodes a polypeptide comprising at least 6

CC consecutive amino acids of any of the 55 250-500 residue amino acid  
CC sequences (S1), given in the specification, altering the environmental  
CC stress response or tolerance of a plant, or altering a plant's trait and  
CC altering the expression levels of at least one gene in a plant. The  
CC recombinant polynucleotide and methods are useful for altering the  
CC regulation of gene expression of plants to modify the plant's traits, in  
CC particular with respect to environmental stress responses (e.g. to viral  
CC infection, fungal infection, microbial infection, herbicide resistance,  
CC heat, cold, heavy metals, low light, drought, osmotic stress and salt  
CC concentration). The present sequence is an environmental stress related  
CC polynucleotide of the invention.

CC Sequence 1432 BP; 426 A; 269 C; 351 G; 386 T; 0 U; 0 Other;

Query Match 9.7%; Score 129.2; DB 10; Length 1432;

Best Local Similarity 70.3%; Pred. No. 8e-12; Mismatches 73; Indels 0; Gaps 0;

Matches 173; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 413 TCCAAGCGGAGCAGAGAGCGCGCATCCCATGACGAGAGAGACAGAGCTG 472  
DB 580 TCMAAGTCGATCAAGAGAGAGAGGTATCGGTGACAGAGATGAGACAGTTA 639  
QY 473 TTCTTGCTGGGCTGAGCAAGTTCCGCAAGGGGACTGGCGGACATCTCCGCACTTC 532  
DB 640 TTTCTTCTTGTTGGATTAAGTACGCGAAAGGTGATGGCGTATTTCTCGCAACTTT 699  
QY 533 GTCACTCTGGCGGACCGCAACGAGTGGCGAGCGGACGAGAGAACTACTTCTTCGCGCTC 592  
DB 700 GTAGTAACAGAGAACCGACCCCAAGTTGCGAGCCATGCTCAAAAGTATTTCAATTGCTTA 759  
QY 593 AACTCATGAACCGGACCGCGCGCTCCAGCATTCACGACATCAGCAGCGCTCACCGCC 652  
DB 760 AATTCAATGAACAAAGACAGAGAGGATCAAGCATTCACGACATCACTAAGTGTGGCAAC 819  
QY 653 GCGCAT 658  
DB 820 GCAGAT 825

RESULT 7  
AD161458  
ID AD161458 standard; cDNA; 1432 BP.

XX AD161458;

DT 22-APR-2004 (first entry)

DE cDNA encoding A. thaliana novel transcription factor (TF) #79.

KW Plant; transcription factor; TF; plant trait; gene; se.

OS Arabidopsis thaliana.

PN US2003229915-A1.

PD 11-DEC-2003.

PF 22-NOV-2002; 2002US-00302267.

PR 18-FEB-1999; 99US-0120880P.  
PR 22-FEB-1999; 99US-0121037P.  
PR 11-MAR-1999; 99US-0124276P.  
PR 15-APR-1999; 99US-0128450P.  
PR 20-MAY-1999; 99US-0135134P.  
PR 15-JUL-1999; 99US-0144153P.  
PR 22-OCT-1999; 99US-0161143P.  
PR 01-NOV-1999; 99US-0162656P.  
PR 17-FEB-2000; 2000US-00506720.

PA (KEDD/) KEDDIE J.  
PA (FROM/) FROMM M.  
PA (HEAR/) HEARD J.  
PA (RIEC/) RIECHMANN J L.

PA (ADAM/) ADAM L.  
PA (BROU/) BROUN P.  
PA (PINE/) PINEDA O.  
PA (REUB/) REUBER L.  
PA (ZHAN/) ZHANG J.  
PA (YUOG/) YU G.  
PA (JIANG/) JIANG C.  
PA (SAMA/) SAMAHA R.  
PA (PILG/) PILGRIM M.  
PA (CREB/) CREBELMAN R.

PI Keddle J, Fromm M, Heard J, Riechmann JL, Adam L, Broun P,  
PI Pineda O, Reuber L, Zhang J, Yu G, Jiang C, Samaha R, Pilgrim M;  
PI Creelman R;

DR WPI: 2004-052052/05.  
DR P-PSDB; AD161459.

PT New polynucleotide, useful in screening for a transcription factor that  
PT modifies a plant trait.

PS Claim 1; SEQ ID NO 157; 16pp; English.

CC The present invention relates to the isolation of novel plant  
CC (arabidopsis thaliana) polynucleotide sequences that encode transcription  
CC factors (TFs), and the polypeptide sequences for the TFs. The  
CC polynucleotide sequences are useful in screening for a transcription  
CC factor that modifies a plant trait. Also disclosed is an expression  
CC vector comprising a TF polynucleotide sequence, a host cell comprising  
CC the expression vector, a transgenic plant comprising or ectopically  
CC expressing an isolated TF polynucleotide sequence, a method for screening  
CC for a molecule that modifies a plant trait, a method for producing a TF  
CC transgenic plant, a method for identifying a sequence homologous to a TF  
CC polynucleotide or polypeptide sequence, and a method for screening for a  
CC transcription factor that modifies a plant trait. The present sequence  
CC encodes a novel A. thaliana TF of the invention. Note: The sequence data  
CC for this patent did not form part of the printed specification. The  
CC complete sequence data for this patent was obtained in electronic format  
CC directly from the USPTO web site at seqdata.uspto.gov.

XX SQ Sequence 1432 BP; 426 A; 269 C; 351 G; 386 T; 0 U; 0 Other;

Query Match 9.7%; Score 129.2; DB 12; Length 1432;

Best Local Similarity 70.3%; Pred. No. 8e-12; Mismatches 73; Indels 0; Gaps 0;

Matches 173; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 413 TCCAAGCGGAGCAGAGAGCGCGCATCCCATGACGAGAGAGACAGAGCTG 472  
DB 580 TCMAAGTCGATCAAGAGAGAGAGGTATCGGTGACAGAGATGAGACAGAGTTA 639  
QY 473 TTCTTGCTGGGCTGAGCAAGTTCCGCAAGGGGACTGGCGGACATCTCCGCACTTC 532  
DB 640 TTTCTTCTTGTTGGATTAAGTACGCGAAAGGTGATGGCGTATTTCTCGCAACTTT 699  
QY 533 GTCACTCTGGCGGACCGCAACGAGTGGCGAGCGGACGAGAGAACTACTTCTTCGCGCTC 592  
DB 700 GTAGTAACAGAGAACCGACCCCAAGTTGCGAGCCATGCTCAAAAGTATTTCAATTGCTTA 759  
QY 593 AACTCATGAACCGGACCGCGCGCTCCAGCATTCACGACATCAGCAGCGCTCACCGCC 652  
DB 760 AATTCAATGAACAAAGACAGAGAGGATCAAGCATTCACGACATCACTAAGTGTGGCAAC 819  
QY 653 GCGCAT 658  
DB 820 GCAGAT 825

RESULT 8

AD002450  
ID AD002450 standard; cDNA; 1432 BP.

XX AD002450;

DT 01-JUL-2004 (first entry)  
XX  
XX Thalecress transcription factor CDNA #432.  
XX Thalecress; transcription factor; ss; gene; plant; transgenic;  
KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;  
KW phosphate limitation; potassium limitation; nitrogen limitation;  
KW hormone sensitivity; disease resistance; sugar sensing; seed germination;  
KW flowering; inflorescence architectural change;  
KW meristem cell differentiation; phyllotaxy; apical dominance;  
KW trichome development; seed development; premature senescence;  
KW delayed senescence; lethality; necrosis; plant size; leaf morphology;  
KW seed morphology; secondary metabolism; light response; shade avoidance.  
XX  
XX Arabidopsis thaliana.  
XX  
XX US2004045049-A1.  
XX  
XX 04-MAR-2004.  
XX  
XX 10-APR-2003; 2003US-00412699.  
XX  
XX 13-SEP-1999; 99US-00394519.  
XX 21-FEB-2000; 2000US-00489376.  
XX 17-FEB-2000; 2000US-00506720.  
XX 22-MAR-2000; 2000US-00532591.  
XX 22-MAR-2000; 2000US-00533029.  
XX 22-MAR-2000; 2000US-00533030.  
XX 22-MAR-2000; 2000US-00533392.  
XX 22-MAR-2000; 2000US-00533648.  
XX 06-APR-2000; 2000WO-00009448.  
XX 16-NOV-2000; 2000US-00713994.  
XX 27-MAR-2001; 2001US-00819142.  
XX 17-APR-2001; 2001US-00837444.  
XX 30-JAN-2002; 2002US-00958131.  
XX 14-JUN-2002; 2002US-00171468.  
XX 09-AUG-2002; 2002US-00225066.  
XX 09-AUG-2002; 2002US-00225067.  
XX 09-AUG-2002; 2002US-00225068.  
XX 17-DEC-2002; 2002US-0434166P.  
XX 25-FEB-2003; 2003US-00374780.  
XX  
XX (ZHAN/) ZHANG J.  
XX (FROM/) FROM M E.  
XX (HEAR/) HEARD J E.  
XX (RIEC/) RIECHMANN J L.  
XX (ADAM/) ADAM L J.  
XX (BROU/) BROUN P E.  
XX (PINE/) PINEDA O.  
XX (REUB/) REUBER T L.  
XX (KEDD/) KEDDIE J S.  
XX (YUGG/) YU G.  
XX (JIAN/) JIANG C.  
XX (SAMA/) SAMAH R S.  
XX (PILG/) PILGRIM M L.  
XX (CREE/) CREELMAN R A.  
XX (DUBE/) DUBBEL A N.  
XX (RATC/) RATCLIFFE O.  
XX (KUMI/) KUMIMOTO R.  
XX (SHER/) SHERMAN B K.  
XX  
XX Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;  
PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;  
PI Pilgrim M, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;  
PI Sherman B;  
XX  
XX WPI; 2004-225755/21.  
XX P-PSDB; ADO02451.  
XX  
XX New transgenic plant, useful in developing phenotypes with altered or  
PT improved characteristics or traits.  
XX  
XX Claim 1; SEQ ID NO 863; 213pp; English.

XX The invention relates to a transgenic plant comprises a recombinant  
CC polynucleotide having a polynucleotide sequence or its complementary  
CC sequence comprising a sequence encoding a polypeptide, that initiates  
CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,  
CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588  
CC ADO05527 or ADO03530-ADO03559. Also included are using a transgenic  
CC plant to grow a progeny plant, an expression cassette (comprising a  
CC constitutive, inducible or tissue-specific promoter and a recombinant  
CC polynucleotide described above), a host cell comprising the expression  
CC cassette, producing a modified plant having a modified trait, identifying  
CC a factor that is modulated by or interacts with a polypeptide encoded by  
CC the polynucleotide sequence and identifying at least one downstream  
CC polynucleotide sequence that is subject to a regulatory effect of any of  
CC the polypeptides encoded by the polynucleotide described above. The  
CC transgenic plant is useful for producing a plant that has an altered  
CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance  
CC to chilling, germination in cold conditions freezing tolerance, tolerance  
CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to  
CC salt, tolerance to phosphate limitation, tolerance to potassium  
CC limitation, decreased sensitivity to nitrogen limitation), altered  
CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered  
CC response to ethylene, disease resistance, altered susceptibility to  
CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to  
CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered  
CC susceptibility to Sclerotinia, altered sugar sensing, improved seed  
CC germination and seedling vigor, early flowering, late flowering, extended  
CC period of flowering, an inflorescence architectural change, a change in  
CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell  
CC differentiation, altered phyllotaxy, altered branching pattern, reduced  
CC apical dominance, reduced trichome density, ectopic trichome development,  
CC altered trichome development, altered stem morphology, increased root  
CC growth, increased root hairs, altered seed development, altered cell  
CC proliferation/cell differentiation, premature senescence, delayed  
CC senescence, lethality, increased necrosis, an increase in seedling or  
CC plant size, decreased plant size, a change in leaf morphology, increased  
CC altered leaf development, increased leaf size and mass, glossy leaves,  
CC leaf cell expansion, change in seed morphology, altered seed coloration,  
CC increased seed size, decreased seed size, altered seed shape, change in  
CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid  
CC content, increased leaf insoluble sugars, decreased leaf insoluble  
CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid  
CC content, an alteration of leaf glucosinolate content, change in seed  
CC biochemistry, an increase in seed oil content, decrease in seed oil  
CC content, increase in seed fatty acid content, decrease in seed fatty acid  
CC content, increase in seed protein content, decrease in seed protein  
CC content, alteration in seed prenyl lipid content, increase in seed  
CC sterols, upregulation of genes involved in secondary metabolism, increase  
CC in root anthocyanins, increase in plant anthocyanins, and alteration in  
CC light response or shade avoidance. The present sequence encodes a  
CC thalecress transcription factor of the invention.  
XX  
XX Sequence 1432 BP; 426 A; 269 C; 351 G; 386 T; 0 U; 0 Other;  
SQ  
Query Match 9.7%; Score 129.2; DB 12; Length 1432;  
Best Local Similarity 70.3%; Pred. No. 8e-12;  
Matches 173; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
Qy 413 TCCAGGCGAGAGAGAGAGAGAGGCGATCCATGACGAGAGAGACAGGCTG 472  
Db 580 TCAAGTCCGATCAAGACACGAAAGGATGCGGTGACAGAAATGACACAGTTA 639  
Qy 473 TTCTGCTGGGGGTGACAAATTGCGACAGGGGAGTGGCGAGCATCTTGGCCAACTTC 532  
Db 640 TTCTTCTTGTGTGATGATGTAAGGGAAGTGATGGCGTGAATTTCTCCCACTTT 639  
Qy 533 GTCATCTCGGAGCGCCAAAGCGAGTGGCGACGCGCGAGAAAGTACTTCGCGCTC 592  
Db 700 GTAGTAACAAGAACACCGACCAAGTTGCGAGCGCATGCTCAAAAGTATTCATTCGCTA 759  
Qy 593 AACTCGATGAACCGCGCGCGCGCTCCAGATCCAGACAGCATCAACAGCGCC 652  
Db 760 AATTCAATGAACAAGACAGAGAGCGCATCAAGCATTCACGATCACTAGTTGGCAAC 819



[illegible]

RESULT 11	
ADQ03867	
ID	ADQ03867 standard; DNA; 257 BP.
XX	
AC	
XX	ADQ03867;
XX	
DT	09-SEP-2004 (first entry)
XX	
DE	Maize transcription factor segid 713.
XX	
KW	maize; soybean; Arabidopsis thaliana; transcription factor; homeobox
KW	HLH; leucine zipper; zinc finger; transformed plant; plant;
KW	metabolic pathway; mutation detection; polymorphism; plant trait;
KW	genome mapping; gene identification; gene analysis; plant breeding;
KW	transgenic; ds.
XX	
OS	Zea mays.
XX	
PN	US2004123339-A1.
XX	
PD	24-JUN-2004.
XX	
PF	06-AUG-2001; 2001US-00922293.
XX	
PR	24-NOV-1997; 97US-0065700P.
PR	09-DEC-1997; 97US-0065947P.
PR	13-JAN-1998; 98US-0071479P.
PR	10-FEB-1998; 98US-0074201P.
PR	10-FEB-1998; 98US-0074280P.
PR	10-FEB-1998; 98US-0074281P.
PR	10-FEB-1998; 98US-0074282P.
PR	12-FEB-1998; 98US-0074565P.
PR	12-FEB-1998; 98US-0074566P.
PR	12-FEB-1998; 98US-0074567P.
PR	19-FEB-1998; 98US-0074789P.
PR	19-FEB-1998; 98US-0075459P.
PR	19-FEB-1998; 98US-0075460P.
PR	19-FEB-1998; 98US-0075461P.
PR	19-FEB-1998; 98US-0075462P.
PR	19-FEB-1998; 98US-0075463P.
PR	19-FEB-1998; 98US-0075464P.
PR	09-MAR-1998; 98US-0077229P.
PR	09-MAR-1998; 98US-0077230P.
PR	09-MAR-1998; 98US-0077231P.
PR	18-MAR-1998; 98US-0078366P.
PR	07-APR-1998; 98US-0080844P.
PR	27-APR-1998; 98US-0083067P.
PR	29-APR-1998; 98US-0083366P.

PR	29-APR-1998;	98US-0083387P.
PR	29-APR-1998;	98US-0083388P.
PR	29-APR-1998;	98US-0083389P.
PR	13-MAY-1998;	98US-0085222P.
PR	13-MAY-1998;	98US-0085223P.
PR	13-MAY-1998;	98US-0085224P.
PR	15-MAY-1998;	98US-0085533P.
PR	21-MAY-1998;	98US-0086183P.
PR	21-MAY-1998;	98US-0086184P.
PR	21-MAY-1998;	98US-0086185P.
PR	21-MAY-1998;	98US-0086186P.
PR	21-MAY-1998;	98US-0086187P.
PR	21-MAY-1998;	98US-0086188P.
PR	16-JUN-1998;	98US-0089524P.
PR	16-JUN-1998;	98US-0089793P.
PR	18-JUN-1998;	98US-0089806P.
PR	18-JUN-1998;	98US-0089807P.
PR	18-JUN-1998;	98US-0089808P.
PR	18-JUN-1998;	98US-0089810P.
PR	18-JUN-1998;	98US-0089811P.
PR	18-JUN-1998;	98US-0089812P.
PR	18-JUN-1998;	98US-0089813P.
PR	18-JUN-1998;	98US-0089814P.
PR	30-JUN-1998;	98US-0091247P.
PR	30-JUN-1998;	98US-0091405P.
PR	09-SEP-1998;	98US-0093667P.
PR	09-SEP-1998;	98US-0093668P.
PR	09-SEP-1998;	98US-0093670P.
PR	09-SEP-1998;	98US-0093697P.
PR	16-SEP-1998;	98US-0100672P.
PR	16-SEP-1998;	98US-0100673P.
PR	16-SEP-1998;	98US-0100674P.
PR	17-SEP-1998;	98US-0100963P.
PR	21-SEP-1998;	98US-0101130P.
PR	21-SEP-1998;	98US-0101131P.
PR	21-SEP-1998;	98US-0101132P.
PR	22-SEP-1998;	98US-0101343P.
PR	22-SEP-1998;	98US-0101344P.
PR	22-SEP-1998;	98US-0101347P.
PR	22-SEP-1998;	98US-0101508P.
PR	25-SEP-1998;	98US-0101707P.
PR	13-OCT-1998;	98US-0104124P.
PR	13-OCT-1998;	98US-0104126P.
PR	13-OCT-1998;	98US-0104127P.
PR	13-OCT-1998;	98US-0104128P.
PR	18-NOV-1998;	98US-0106996P.
PR	19-NOV-1998;	98US-0109018P.
PR	24-NOV-1998;	98US-0019129P.
PR	08-DEC-1998;	98US-0021029P.
PR	11-DEC-1998;	98US-0111981P.
PR	22-DEC-1998;	98US-0113224P.
PR	12-JAN-1999;	99US-00229413.
XX	(CONN/) CONNER T W.	
PA	(HECK/) HECK G R.	
PA	(LIU/) LIU J.	
XX		
PI	Conner TW, Heck GR, Liu J;	
XX		
DR	WPI; 2004-468202/44.	
XX		
PT	New substantially purified nucleic acid molecule that encodes a maize, soybean or Arabidopsis thaliana transcription factor or its fragment, useful for genome mapping, gene identification and analysis or plant breeding.	
PT		
XX		
PS	Claim 2; SEQ ID NO 713; 140bp; English.	
XX		
CC	The invention describes a substantially purified nucleic acid molecule that encodes a maize, soybean or Arabidopsis thaliana transcription factor or its fragment, where the maize or soybean transcription factor is homeobox, HbH, leucine zipper, zinc finger, or other transcription factor. Also described are: a substantially purified maize or soybean	
CC		





PA (L1UJ/) Liu J.  
XX  
PI Conner TW, Heck GR, Liu J;  
XX  
XX WPI; 2004-468202/44.  
XX  
PT New substantially purified nucleic acid molecule that encodes a maize,  
PT soybean or Arabidopsis thaliana transcription factor or its fragment,  
PT useful for genome mapping, gene identification and analysis or plant  
breeding.  
PS Claim 2; SEQ ID NO 1242; 140pp; English.  
XX  
CC The invention describes a substantially purified nucleic acid molecule  
CC that encodes a maize, soybean or Arabidopsis thaliana transcription  
CC factor or its fragment, where the maize or soybean transcription factor  
CC is homeobox, HLH, leucine zipper, zinc finger, or other transcription  
CC factor. Also described are a substantially purified maize or soybean  
CC transcription factor or its fragment defined above; a substantially  
CC purified antibody or its fragment which is capable of specifically  
CC binding to the transcription factor or its fragment above; a transformed  
CC plant; a method for determining a level or pattern in a plant cell of a  
CC transcription factor in a plant metabolic pathway; a method of  
CC determining a mutation in a plant; a method of producing a plant  
CC containing an overexpressed protein or reduced levels of plant  
CC transcription factor; a method of determining an association between a  
CC polymorphism and a plant trait; and a method of isolating a nucleic acid  
CC that encodes a plant transcription factor or its fragment. The nucleic  
CC acid molecules, proteins and their fragments are useful for genome  
CC mapping, gene identification and analysis, plant breeding, preparation of  
CC constructs for use in plant gene expression and transgenic plants. The  
CC nucleic acid molecules are useful as markers or probes. This sequence  
CC represents a maize transcription factor polynucleotide.  
XX  
SQ Sequence 386 BP; 81 A; 130 C; 120 G; 55 T; 0 U; 0 Other;  
  
Query Match 9.0%; Score 120.2; DB 12; Length 386;  
Best Local Similarity 77.2%; Pred. No. 2.3e-10;  
Matches 146; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
  
QY 461 GAGCAGAGGCTGTCTTCTGCTGGGCTGGACAAGTTCGCGAAGGGAAGTGGCGAGATC 520  
Db |||||  
62 GACAAAGGCTGTTCCATTAGCTGAAAGATATCGCGAAGGGGACTGAGAGATTA 121  
QY 521 TCGCGCAACTTGTCTATCTCGCGAAGCGCAAGAGTGGGAGCCAGCGCAGAGTAC 580  
Db |||||  
122 TCGCGCAACTAGCTTCAGACCGGACGCCCAAGGAGTGGCGAAGCGCAGAGTAC 181  
QY 581 TTATCGGCGCTCACTCATTAACCGGACCGCGCGCTCCAGCATTCAGAGATCACC 640  
Db |||||  
182 TTATCGGCGCTCACTCGCGGCGAAGGACAAAGAGAGTTCACATTCAGATCACC 241  
QY 641 AGCGTCACC 649  
Db |||||  
242 ACGGTGAAC 250  
  
RESULT 13  
ADQ03772  
ID ADQ03772 standard; DNA; 315 BP.  
XX  
AC ADQ03772;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Maize transcription factor seqid 618.  
XX  
KM maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;  
KM HLH; leucine zipper; zinc finger; transformed plant; plant;  
KM metabolic pathway; mutation detection; polymorphism; plant trait;  
KM genome mapping; gene identification; gene analysis; plant breeding;  
KM transgenic; ds.  
XX

OS Zea mays.  
XX  
PN US2004123339-A1.  
XX  
PD 24-JUN-2004.  
XX  
PF 06-AUG-2001; 2001US-00922293.  
XX  
XX 24-NOV-1997; 97US-0067000P.  
PR 09-DEC-1997; 97US-0069472P.  
PR 13-JAN-1998; 98US-0071479P.  
PR 10-FEB-1998; 98US-0074201P.  
PR 10-FEB-1998; 98US-0074280P.  
PR 10-FEB-1998; 98US-0074281P.  
PR 10-FEB-1998; 98US-0074282P.  
PR 12-FEB-1998; 98US-0074565P.  
PR 12-FEB-1998; 98US-0074566P.  
PR 12-FEB-1998; 98US-0074567P.  
PR 19-FEB-1998; 98US-0074789P.  
PR 19-FEB-1998; 98US-0075459P.  
PR 19-FEB-1998; 98US-0075460P.  
PR 19-FEB-1998; 98US-0075461P.  
PR 19-FEB-1998; 98US-0075462P.  
PR 19-FEB-1998; 98US-0075463P.  
PR 19-FEB-1998; 98US-0075464P.  
PR 09-MAR-1998; 98US-0077229P.  
PR 09-MAR-1998; 98US-0077230P.  
PR 09-MAR-1998; 98US-0077231P.  
PR 18-MAR-1998; 98US-0078368P.  
PR 07-APR-1998; 98US-0080844P.  
PR 27-APR-1998; 98US-0083067P.  
PR 29-APR-1998; 98US-0083386P.  
PR 29-APR-1998; 98US-0083387P.  
PR 29-APR-1998; 98US-0083388P.  
PR 29-APR-1998; 98US-0083389P.  
PR 13-MAY-1998; 98US-0085222P.  
PR 13-MAY-1998; 98US-0085223P.  
PR 13-MAY-1998; 98US-0085224P.  
PR 15-MAY-1998; 98US-0085533P.  
PR 21-MAY-1998; 98US-0086183P.  
PR 21-MAY-1998; 98US-0086184P.  
PR 21-MAY-1998; 98US-0086185P.  
PR 21-MAY-1998; 98US-0086186P.  
PR 21-MAY-1998; 98US-0086187P.  
PR 21-MAY-1998; 98US-0086188P.  
PR 16-JUN-1998; 98US-0089524P.  
PR 18-JUN-1998; 98US-0089733P.  
PR 18-JUN-1998; 98US-0089806P.  
PR 18-JUN-1998; 98US-0089807P.  
PR 18-JUN-1998; 98US-0089808P.  
PR 18-JUN-1998; 98US-0089810P.  
PR 18-JUN-1998; 98US-0089811P.  
PR 18-JUN-1998; 98US-0089812P.  
PR 18-JUN-1998; 98US-0089813P.  
PR 18-JUN-1998; 98US-0089814P.  
PR 30-JUN-1998; 98US-0091247P.  
PR 30-JUN-1998; 98US-0091405P.  
PR 09-SEP-1998; 98US-0096677P.  
PR 09-SEP-1998; 98US-0096680P.  
PR 09-SEP-1998; 98US-0096681P.  
PR 09-SEP-1998; 98US-0096682P.  
PR 09-SEP-1998; 98US-0096683P.  
PR 16-SEP-1998; 98US-0100672P.  
PR 16-SEP-1998; 98US-0100673P.  
PR 16-SEP-1998; 98US-0100674P.  
PR 17-SEP-1998; 98US-0100675P.  
PR 21-SEP-1998; 98US-0101130P.  
PR 21-SEP-1998; 98US-0101131P.  
PR 21-SEP-1998; 98US-0101132P.  
PR 22-SEP-1998; 98US-0101343P.  
PR 22-SEP-1998; 98US-0101344P.  
PR 22-SEP-1998; 98US-0101347P.  
PR 22-SEP-1998; 98US-0101508P.  
PR 25-SEP-1998; 98US-0101707P.



PR 09-SEP-1998; 98US-0099670P.  
PR 09-SEP-1998; 98US-0099671P.  
PR 16-SEP-1998; 98US-0100672P.  
PR 16-SEP-1998; 98US-0100673P.  
PR 16-SEP-1998; 98US-0100674P.  
PR 17-SEP-1998; 98US-0100963P.  
PR 21-SEP-1998; 98US-0101131P.  
PR 21-SEP-1998; 98US-0101132P.  
PR 21-SEP-1998; 98US-0101133P.  
PR 22-SEP-1998; 98US-0101343P.  
PR 22-SEP-1998; 98US-0101344P.  
PR 22-SEP-1998; 98US-0101347P.  
PR 22-SEP-1998; 98US-0101508P.  
PR 23-SEP-1998; 98US-0101707P.  
PR 13-OCT-1998; 98US-0104124P.  
PR 13-OCT-1998; 98US-0104126P.  
PR 13-OCT-1998; 98US-0104127P.  
PR 13-OCT-1998; 98US-0104128P.  
PR 18-NOV-1998; 98US-0108996P.  
PR 19-NOV-1998; 98US-0109018P.  
PR 24-NOV-1998; 98US-0019912P.  
PR 08-DEC-1998; 98US-00210297.  
PR 11-DEC-1998; 98US-0111981P.  
PR 22-DEC-1998; 98US-0113224P.  
PR 12-JAN-1999; 99US-00229413.

XX (CONN/) CONNER T W.  
PA (HECK/) HECK G R.  
PA (LIU/) LIU J.

XX Conner TW, Heck GR, Liu J;

XX WPI; 2004-466202/44.

PT New substantially purified nucleic acid molecule that encodes a maize,  
PT soybean or Arabidopsis thaliana transcription factor or its fragment,  
PT useful for genome mapping, gene identification and analysis or plant  
PT breeding.

XX Claim 2; SEQ ID NO 707; 140bp; English.

XX The invention describes a substantially purified nucleic acid molecule  
CC that encodes a maize, soybean or Arabidopsis thaliana transcription  
CC factor or its fragment, where the maize or soybean transcription factor  
CC is homeobox, HLF, leucine zipper, zinc finger, or other transcription  
CC factor. Also described are: a substantially purified maize or soybean  
CC transcription factor or its fragment defined above; a substantially  
CC purified antibody or its fragment which is capable of specifically  
CC binding to the transcription factor or its fragment above; a transformed  
CC plant; a method for determining a level or pattern in a plant cell of a  
CC transcription factor in a plant metabolic pathway; a method of  
CC determining a mutation in a plant; a method of producing a plant  
CC containing an overexpressed protein or reduced levels of plant  
CC transcription factor; a method of determining an association between a  
CC polymorphism and a plant trait; and a method of isolating a nucleic acid  
CC that encodes a plant transcription factor or its fragment. The nucleic acid  
CC molecules, proteins and their fragments are useful for genome  
CC mapping, gene identification and analysis, plant breeding, preparation of  
CC constructs for use in plant gene expression and transgenic plants. The  
CC nucleic acid molecules are useful as markers or probes. This sequence  
CC represents a maize transcription factor polynucleotide.

XX Sequence 144 BP; 27 A; 39 C; 54 G; 24 T; 0 U; 0 Other;

SQ Query Match 8.9%; Score 119; DB 12; Length 144;

Best Local Similarity 89.5%; Pred. No. 3.6e-10;

Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 428 GAGAGGCGCAAGGGGATCCATGACGAGGAGAGAGCAAGGCTGTTCTTGCTGGGCTG 487  
DB 2 GAGGCGGCGCAAGGGGCTGCTGACGAGAGAGAGCAAGGCTGTTCTTGCTGGATCTG 61  
QY 488 GACAAGTTGGCAAGGGGAGCTGGCGAGCATCTTCGGCAACTTCGTACTCTCGGAGC 547

DB 62 GACAAGTTCTGCAAGGGCGAGCTGCGAGCATCTCGCAACTTCGTACTCTCGGAGC 121

QY 548 CCAACGAGGTGGCGAGCCACGC 570  
DB 122 CCGACGAGGTGGCGAGCCACGC 144

RESULT 15  
ABK65275  
ID ABK65275 standard; cDNA; 867 BP.

AC ABK65275;

DT 02-JUL-2002 (first entry)

XX Arabidopsis cDNA encoding a transcription factor #127.

KM Plant, sex; gene; transcription factor; transgenic; agriculture;  
KM metabolic chemical; environmental stress; drought;  
KM microbial disease resistance; herbicide resistance; seed yield;  
KM fruit yield; growth rate; leaf senescence; flower senescence.

OS Arabidopsis thaliana.

XX WO200215675-A1.

XX 28-FEB-2002.

PF 22-AUG-2001; 2001WO-US026189.

XX 22-AUG-2000; 2000US-0227439P.

PR 16-NOV-2000; 2000US-00713994.

PR 18-APR-2001; 2001US-00837944.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (PIUG/) PILGRIM M.

PA (CREE/) CREELMAN R.

PA (DUBE/) DUBEL J.

PA (HEAR/) HEARD J.

PA (JIANG/) JIANG C.

PA (KEDD/) KEDDIE J.

PA (ADAM/) ADAM L.

PA (RATC/) RATCLIFF O.

PA (REUB/) REUBER J L.

PA (RIEC/) RIECHMANN J L.

PA (YUGG/) YU G.

PA (PINE/) PINEDA O.

XX Pilgrim M, Creelman R, Dubel J, Heard J, Jiang C, Keddie J,

PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;

XX WPI; 2002-292022/33.

DR P-PSDB; AAU93089.

XX An isolated or recombinant polynucleotide used to produce a transgenic

CC plant.

XX Claim 4; Page 527-528; 941bp; English.

CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
CC encoding an Arabidopsis thaliana transcription factor, their variants,  
CC complements, fragments, or related polynucleotide with 31% to 95%

CC sequence identity, where the plant possesses an altered trait as compared  
CC to a wild-type or reference plant, or the plant exhibits an altered

CC phenotype as compared to a wild-type or reference plant, or the plant  
CC exhibits ectopic expression or altered expression of one or more genes

CC associated with a plant trait as compared to a wild plant. Also included  
CC are a transgenic plant comprising the polynucleotides, a computer

CC readable medium having stored sequence information, and identifying a  
CC homologue sequence from a database comprising a plurality of known plant

CC sequences comprising inputting sequence information selected from one of  
CC 464 fully defined sequences given in the specification. The isolated or



**This Page Blank (uspto)**



QY 61 GGGCGCGCGCTGACCAAGGAGCAAGAGCGCTTGAGAGCGCGCTCGCGCTTCGCGC 120  
DB 61 GGGCGCGCGCTGACCAAGGAGCAAGAGCGCTTGAGAGCGCGCTCGCGCTTCGCGC 120  
QY 121 GGGCGCGCGCGCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 180  
DB 121 GGGCGCGCGCGCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 180  
QY 181 GAGCGTCCCGGCGCGAGGTGCGCGAGAGGTGCGAGGAGCACTACGAGCGCGCTGCGA 240  
DB 181 GAGCGTCCCGGCGCGAGGTGCGCGAGAGGTGCGAGGAGCACTACGAGCGCGCTGCGA 240  
QY 241 GAGCGTCCCGGCGCGAGGTGCGCGAGAGGTGCGAGGAGCACTACGAGCGCGCTGCGA 300  
DB 241 GAGCGTCCCGGCGCGAGGTGCGCGAGAGGTGCGAGGAGCACTACGAGCGCGCTGCGA 300  
QY 301 GCGCGCGCGCGCGAGCGAGC 360  
DB 301 GCGCGCGCGCGCGAGCGAGC 360  
QY 361 GCGCGAGAGCGCAAGGC 420  
DB 361 GCGCGAGAGCGCAAGGC 420  
QY 421 GAGAGAGAGAGCGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 480  
DB 421 GAGAGAGAGAGCGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 480  
QY 481 GGGCGTGGAGCAAGTTCGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 540  
DB 481 GGGCGTGGAGCAAGTTCGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 540  
QY 541 GCGGAGCGCAAGCGAGGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 600  
DB 541 GCGGAGCGCAAGCGAGGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 600  
QY 601 GAACCGGAGC 660  
DB 601 GAACCGGAGC 660  
QY 661 GGTGCGCGCGAGAGGC 720  
DB 661 GGTGCGCGCGAGAGGC 720  
QY 721 GGGCGCGCTGGC 780  
DB 721 GGGCGCGCTGGC 780  
QY 781 GC 840  
DB 781 GC 840  
QY 841 GC 900  
DB 841 GC 900  
QY 901 GGTGCGCGCGAGAGGC 960  
DB 901 GGTGCGCGCGAGAGGC 960  
QY 961 GAGAGAGAGAGC 1020  
DB 961 GAGAGAGAGAGC 1020  
QY 1021 GCGGTGC 1080  
DB 1021 GCGGTGC 1080  
QY 1081 GATGAGC 1140  
DB 1081 GATGAGC 1140

QY 1141 GACTGTTGGCAAGTATCATACGTACACCGCGTGAAAGCTTGAATCCAAATGTGATGT 1200  
DB 1141 GACTGTTGGCAAGTATCATACGTACACCGCGTGAAAGCTTGAATCCAAATGTGATGT 1200  
QY 1201 AATTAGCAGCGCGCTTCGTAAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1260  
DB 1201 AATTAGCAGCGCGCTTCGTAAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1260  
QY 1261 TAGGATTTGATTTTCATTTTGTCTGATGTAATTTGCGCACTGTCTCATTTGACCAAAAA 1320  
DB 1261 TAGGATTTGATTTTCATTTTGTCTGATGTAATTTGCGCACTGTCTCATTTGACCAAAAA 1320  
QY 1321 AAAAAAAAAA 1330  
DB 1321 AAAAAAAAAA 1330

RESULT 2  
US-10-437-963-59484/c  
; Sequence 59484, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OR INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ. ID NOS: 204966  
; SEQ. ID NO 59484  
; LENGTH: 1433  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_61102C.1  
US-10-437-963-59484

Query Match 97.3%; Score 1293.8; DB 19; Length 1433;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1298; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ATGACCTCCCAAGGCGCGAG 79  
DB 1433 ATGACCTCCCAAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1374  
QY 80 GAG 139  
DB 1373 GAG 1314  
QY 140 GAG 139  
DB 1313 GAG 1254  
QY 200 TCGGCGAG 259  
DB 1253 TCGGCGAG 1194  
QY 260 GCGGCGCGAG 319  
DB 1193 GCGGCGCGAG 1134  
QY 320 GCGGAG 379  
DB 1133 GCGGAG 1074  
QY 380 GCGGCGCGAG 439



Db	1073	GGCGCGGGGGTAAGACGGGGCGAAGAGCTGTCTCAAGCGGGAGACAGAGAGCGCGAAG	1014
Qy	440	GGCATCCCATGGAACGAGAGAAAGACACAGGCTGTTCTTGCTGGGGCTGGAACAATTCCGC	499
Db	1013	GGCATCCCATGAGCGAGAGAAAGACAGGCGTGTCTTGCTGGGGCTGGAACAAGTTCCGC	954
Qy	500	AAGGGGGGCTGGCGGAGCATCTCGCGCAATTCGTGATCTGGCGGAGCGCAACGAGGTG	559
Db	953	AAGGGGGGCTGGCGGAGCATCTCGCGCAATTCGTGATCTGGCGGAGCGCAACGAGGTG	894
Qy	560	GCGAGCCACGGCGAGAAAGTACTTCACTCCGCTCAATCCATGAACCGGACCGCGCGCG	619
Db	893	GCGAGCCACGGCGAGAAAGTACTTCACTCCGCTCAATCCATGAACCGGACCGCGCGCG	834
Qy	620	TCCAGCATCCACGACATCAACGAGCTACCGCGCGCATCAAGGTGCGCGCGAGAGGGC	679
Db	833	TCCAGCATCCACGACATCAACGAGCTACCGCGCGCATCAAGGTGCGCGCGAGAGGGC	774
Qy	680	GGCCCGATCACCGGCGACCAAGGCGCACCGGGCAACCCCGGGCGGGCGGCTGGAGCCGCGG	739
Db	773	GGCCCGATCACCGGCGACCAAGGCGCACCGGGCAACCCCGGGCGGGCGGCTGGAGCCGCGG	714
Qy	740	GGCATGAAGACCAACACACACCAACCCCGGGCGGGCGCGCGCATGTGAC	799
Db	713	GGCATGAAGACCAACACACACCAACCCCGGGCGGGCGCGCGCATGTGAC	654
Qy	800	AGCGCGGGCCCATGAGGGCACCCGCTGGCGGGCCACATGTGTCCCGCGCGCTGGGACG	859
Db	653	AGCGCGGGCCCATGAGGGCACCCGCTGGCGGGCCACATGTGTCCCGCGCGCTGGGACG	594
Qy	860	CCGGTGTGTGTTCCCGCGGGGCGACGCGCGTACGTCGTGCCGCTGACCCGCGCGCT	919
Db	593	CCGGTGTGTGTTCCCGCGGGGCGACGCGCGTACGTCGTGCCGCTGACCCGCGCGCT	534
Qy	920	CCGGCGAAGATGACCAACATGACGGCGCATGACGAGACATGAGACAGATTTCTTCTCTC	979
Db	533	CCGGCGAAGATGACCAACATGACGGCGCATGACGAGACATGAGACAGATTTCTTCTCTC	474
Qy	980	CTTCTTGATGTACATCTGATTTGTGTTGTGTGAGCGCGGCTACGTCGCTGATC	1039
Db	473	CTTCTTGATGTACATCTGATTTGTGTTGTGTGAGCGCGGCTACGTCGCTGATC	414
Qy	1040	ATCTTGTGTTCTTCTCACAATCTCACTAATGTAAACATACATAGATCAGATGCGAAGATGC	1099
Db	413	ATCTTGTGTTCTTCTCACAATCTCACTAATGTAAACATACATAGATCAGATGCGAAGATGC	354
Qy	1100	AGGGATTGGGGATTAAAGCGCAATAAGTAAGTATTTTCTGACCTGTTTGCAGATGATCA	1155
Db	353	AGGGATTGGGGATTAAAGCGCAATAAGTAAGTATTTTCTGACCTGTTTGCAGATGATCA	294
Qy	1160	TCACGTACACCCGGGAAAGCTTAGCTCCAAATGTGATGTATAATGACGAGCGGCTTCCG	1219
Db	293	TCACGTACACCCGGGAAAGCTTAGCTCCAAATGTGATGTATAATGACGAGCGGCTTCCG	234
Qy	1220	TACGTGTGCGCGCGATGATGATCTTGCAGGGGTTCGAATTTAGGATGATTTTCATTT	1279
Db	233	TACGTGTGCGCGCGATGATGATCTTGCAGGGGTTCGAATTTAGGATGATTTTCATTT	174
Qy	1280	TGCTGATGTAAATTTTGCACATGTCTCATTTGACCAAAAAA	1324
Db	173	TGCTGATGTAAATTTTGCCTACTGCTTCATTTGACCAAAAAACATATA	129

RESULT 3  
 US-10-630-636-4  
 : Sequence 4, Application US/10630636  
 : Publication No. US20040107456A1  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: YU, Su-Hsuey  
 :  
 : TITLE OF INVENTION: PLANT MYB GENES  
 :  
 : FILE REFERENCE: 08919-088001  
 :  
 : CURRENT APPLICATION NUMBER: US/10/630,636

```

; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/399,999
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Oryza sativa (japonica cultivar-group)
US-10-630-636-4

```

Query Match	69.0%;	Score 918;	DB 19;	Length 918;
Best Local Similarity	100.0%;	Pred. No. 4.2e-238;		
Matches 918;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	20	ATTGACTCTCCAGACGAGCCAGCAAGCAACACAGCGCGCGCGCGCGCGCTGTGACACAG	79
Db	1	ATGACTCTCCAGACGAGCCAGCAAGCAACACAGCGCGCGCGCGCGCGCTGTGACACAG	60
OY	80	GAGACGACAAAGCGCTTTCGAGAACGCGCTTCGCGCTTGTGCGCGCGCGCGCTCGCCCGCGAC	139
Db	61	GAGACGACAAAGCGCTTTCGAGAACGCGCTTCGCGCTTGTGCGCGCGCGCGCGCTCGCCCGCGAC	120
OY	140	GAGAGCGCGCGCGCGAGACGACTGTGTTCGCGCGCTTCGCGCGAGCGTGTGCGCGCGCGAGG	159
Db	121	GAGAGCGCGCGCGCGAGACGACTGTGTTCGCGCGCTTCGCGCGAGCGTGTGCGCGCGAGG	180
OY	200	TCGCGCGAGAGAGTGTGCGAGGCACTACGAGAGCGCTGTGTGAGAGACGTTCGCGGCATTCGAC	259
Db	181	TCGCGCGAGAGAGTGTGCGAGGCACTACGAGAGCGCTGTGTGAGAGACGTTCGCGGCATTCGAC	240
OY	260	GCGGCGCGCGTTCGCGCTTCGCGCTACGCGCGGAGAGAGTCCGCGCGCGCGCCCGACGGA	319
Db	241	GCGGCGCGCGTTCGCGCTTCGCGCTACGCGCGGAGAGAGTCCGCGCGCGCGCCCGACGGA	300
OY	320	GCGGAGCGCGCGCGCGCGCGCTTCGAGAGACGCGCGGACACCGCGCGCGACGAGCGCAAGGCG	379
Db	301	GCGGAGCGCGCGCGCGCGCGCTTCGAGAGACGCGCGGACACCGCGCGCGACGAGCGCAAGGCG	360
OY	380	GCGGCGCGCGGAGTACGACGCGCGCGCAAGAGTCTCCAAAGCGCGAGACGAGAGAGCGCAAG	439
Db	361	GCGGCGCGCGGAGTACGACGCGCGCGCAAGAGTCTCCAAAGCGCGAGACGAGAGAGCGCAAG	420
OY	440	GGCATCTCCATGAGACGAGAGAAAGACAGCGCTTCTTGTGCTGTGACAAAGTTCGCG	499
Db	421	GGCATCTCCATGAGACGAGAGAAAGACAGCGCTTCTTGTGCTGTGACAAAGTTCGCG	480
OY	500	AAAGGAGGACTGGCGGAGCATCTCGCGCACTTCGTCATCTCGCGACCGCCAAACGCAAGGTG	559
Db	481	AAAGGAGGACTGGCGGAGCATCTCGCGCACTTCGTCATCTCGCGACCGCCAAACGCAAGGTG	540
OY	560	GCGAGCCACGCGCAGAGAGTACTTCATTCGCGCTCAACTCATATGAAACGCGACCGCGCGCG	619
Db	541	GCGAGCCACGCGCAGAGAGTACTTCATTCGCGCTCAACTCATATGAAACGCGACCGCGCGCG	600
OY	620	TTCAGCATCACAAGCATCACAGCGCTCACGCGCGCATCAGGTGCGCGCGACGAGGCG	679
Db	601	TTCAGCATCACAAGCATCACAGCGCTCACGCGCGCATCAGGTGCGCGCGACGAGGCG	660
OY	680	GCCCCGATCACCGGCAACGAGGCAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	739
Db	661	GCCCCGATCACCGGCAACGAGGCAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	720
OY	740	GGCATGAGAGCAACACCAACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	799
Db	721	GGCATGAGAGCAACACCAACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	780
OY	800	AGCG	859
Db	781	AGCG	840
OY	860	CCGGTGGTGTTCGCT	919

Db 841 CCGGTGATGTTCCCGCGGCGCACGCGCGTACGTGTCGCCCTGAGCTACCGGCGCT 900  
QY 920 CCGGCGCAAGATGACCA 937  
Db 901 CCGGCGCAAGATGACCA 918

RESULT 4  
US-10-425-115-168067  
; Sequence 168067, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 168067  
; LENGTH: 1299  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_84858C.1  
US-10-425-115-168067

Query Match 36.0%; Score 479.2; DB 20; Length 1299;  
Best Local Similarity 75.0%; Pred. No. 2.9e-119;  
Matches 769; Conservative 0; Mismatches 193; Indels 64; Gaps 11;

QY 3 GCGAGATCCACACCGGATGACCTCCAGGCGCGGACGACGACGACGCGGCGGCGG 62  
Db 40 GAGAGAGCGCGCCACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 99  
QY 63 CCGGCGGCGTGAACCGAGGAGGACGACGAGGCGTTGAGAAAGCGCTCGCGCTTTCGCG 122  
Db 100 CCGGCGGCGTGAACCGAGGAGGACGACGAGGCGTTGAGAAAGCGGCGTTCGCGCGG 159  
QY 123 CCGCGCGCGCGCGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182  
Db 160 CCGCGCGCGCGCGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 213  
QY 183 GCGGCGCGGCGGAGGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 242  
Db 214 GCGTGC---GGCGCGAAGCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 270  
QY 243 ACCTCGGCGCATGACGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 299  
Db 271 ACCTCGGCGCATGACGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 330  
QY 300 CCGGCGGCGCGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 350  
Db 331 CCGGCGGAGGCGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGAG 390  
QY 351 -----GCGGACCGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 404  
Db 391 GAGGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 447  
QY 405 AGAGCTCTCCAGGCGGAGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 464  
Db 448 AGAGCTCTCCAGGCGGAGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 507  
QY 465 ACAGGCTGTTCTTCTGCGGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 524  
Db 508 ACAGGCTGTTCTTCTGCGGCGGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 567  
QY 525 GCAACTTGTATCTCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 584  
Db 568 GCAACTTGTATCTCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 627

QY 585 TCCGCTTCAATCTCATGAAACCGGAGCGCGCGCTTCAGATCCAGACATCAACAGCG 644  
Db 628 TCCGCTTCAATCTCATGAAACCGGAGCGCGCGCTTCAGATCCAGACATCAACAGCG 687  
QY 645 TCAACCGCGCGGATGAGTGGCGCGGAGGAGGCGCGCGGATGAGTGGCGCGGAGG 704  
Db 688 TCAACCGCGCGGATGAGTGGCGCGGAGGAGGCGCGCGGATGAGTGGCGCGGAGG 747  
QY 705 CCGGCAACCCCGGCGGCGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 764  
Db 748 CCGG-----GGGATTCGATGAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 792  
QY 765 ACCCGGCG 824  
Db 793 -----CGGTCCCG 840  
QY 825 TCCCG 880  
Db 841 TCCCG 900  
QY 881 --CACGCGCGGATGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 938  
Db 901 ATCACTGCG 960  
QY 939 GACGCGCGATGAGCGG-ACATGAGCAGCATTTCTCTCTCTCTCTCTCTCTCTCT 997  
Db 961 GATCGTCTAGCAGCGCGCATTTGCGCATTTGCGCGCGCGCGCGCGCGCGCGCGCG 1020  
QY 998 TGATTT 1003  
Db 1021 TGATTT 1026

RESULT 5  
US-10-739-930-2650  
; Sequence 2650, Application US/10739930  
; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 2650  
; LENGTH: 1516  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER42\_46  
US-10-739-930-2650

Query Match 33.5%; Score 445.4; DB 20; Length 1516;  
Best Local Similarity 76.8%; Pred. No. 4.3e-110;  
Matches 732; Conservative 1; Mismatches 157; Indels 63; Gaps 13;

QY 28 CCAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 86  
Db 282 CCAAGCATCTGCGATGACCACTTGGCGAGCGGAGGAGGAGGAGGAGGAGGAGG 341  
QY 87 ACAAGCGGTTG-AGAAAGCGGCTCGCGGCTTGGCGGCGCGCGCGCGCGCGCGGAGC 145  
Db 342 ACAAGCGGTTGAGAAAGCGGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGG 395  
QY 146 GCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 205  
Db 396 CCGCGTATGATGAGGCGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 452  
QY 206 GAGGAGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 265  
Db 453 GAGGAGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 512





RESULT 10  
US-10-437-963-47024  
; Sequence 47024, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:

OTHER INFORMATION: CLONE ID: M8145//\_98//5C.1  
US-10-425-115-183322





Query Match 13.6%; Score 180.4; DB 20; Length 1893;  
 Best Local Similarity 60.4%; Pred. No. 2.6e-38;  
 Matches 361; Conservative 0; Mismatches 216; Indels 21; Gaps 3;

Qy	198	GGTCGCGCGAGAGAGTCCGAGCACTACGAGCGCTGTGAGAGACGTCGCGCCATCG	257
Db	495	GGACGGTGGCGAGGTGCTGACGACCTTCAAGCACCTGAGGTGACGTCCAGCAGATCG	554
Qy	258	ACCGGGCCCGGCTCCGCTCCGCGCTACGCGGGAGAGTCCGCGCCGCCGACG	317
Db	555	AGAGCGGGCAGGTGCTTGGCGGCTTACGCGCGGGCCAGCTCGTCACTTCGAGT	614
Qy	318	GAGCGGAGCGCGCGCGCGCGCTCAAGAGCGGGGACACCGCGCGAGAGCGCAAG	377
Db	615	GGAGCGGCTACGCGCCCGCTCGCGG--GGACTTCAAGCACGGGTACCGCTTCCGCGAG	671
Qy	378	GCGCGCGCGCGGAGTACGAGCGGCAAGAGCTGCTCAAGGCGGAGCAGAGAGCGCA	437
Db	672	GCTCGGGAGCGGCGCACCGCGCGACGC-----CGAGCAGAGCGCGAAGA	719
Qy	438	AGGCGATCCATGAGCGAGAGAGACACAGCTGTCTTCTGCGGCTGACAAATTGG	497
Db	720	AGGCGGTGCAATGAGCGAGAGAGACAGGTGTTCTCTAGGCTGAAGAAATACG	779
Qy	498	GCAAGGGGGACTGGCGAGACATCTCGGCAACTTCGTATCTCGGGAACGCCAACGAGG	557
Db	780	GCAAGGTTGACTGAGAGAACATTGCGCAACTTGTCAAGACCAAGACGCCACCAAG	839
Qy	558	TGCGGAGCCACGCGCAGAGTACTTCACTCCGCTCAACTCCATGAAACCGGCGCGCC	617
Db	840	TGGCCAGCCACGCGCAGAGTACTTCACTCCGCTCAACTCCGCGGCAAGAGCAAGAGA	899
Qy	618	GCTCCAGCATCAGCAGCATCACAGCGTC-----ACCGCGGCGATCAGGTGCGCGCGC	671
Db	900	GATCAGCATCAGCAGCATCACAGGTCACAGGTCACAGGTCACAGCAGCAGCGCGCTCGCGT	959
Qy	672	AGCAGGGGCGCGCGCATCACCGGCGCACGAGGCCAACCGCGCGCGCGCGCTGG	731
Db	960	CCCACTCTCTTATATACCAAGCAATGAAAGCAGCAGCTTCAGCTCCGCGCGGGGGA	1019
Qy	732	GCCCGCGGCGATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	789
Db	1020	TATGCCAGTCCCATTTGGCGCGCAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1077

Search completed: July 29, 2005, 22:49:37  
 Job time : 954 secs







	Best Local Similarity	62.2%	Pred. No. 9.3e-12:	
	Matches 196:	Conservative	0: Mismatches 95: Indels 24: Gaps 1	
QY	369	AGCGCAAGGCGCGCGCGCGGGTACGACGCGCGCAAGAGCTCTCCAAAGCGAGCAG	428	
Db	20	AGGGCGCGCGCGGTAGCGGCGAGTCAATCAGAGGAGGAAAGCTCGAGTGGATCAAG	79	
QY	429	AGAGCGCAAGGCGCATCCCATGAGCGAGAGAAAGACAAGCTG-----	472	
Db	80	AGAGAGGAAGGAGATGCTTGAGCCGAGAGAAACACAGGTGGGGTCTTCTTGTGTA	139	
QY	473	-----TTCTTGCTGGGCGCTGACAACTTGGCAAGGGGAGCTGGCGGAGACTCCG	524	
Db	140	ATCGGTTATTTCTTCTAGGCTAGATTAATATGGAAGGCGATTTGCGCAAGTATTTCCC	199	
QY	525	GCAATTTGTATCTTCGGGAGAGCCAAAGCAGGTGGCGAGCCACGCGCAAGTACTTA	584	
Db	200	GGAATTTGTGCMCAAGAGAGCGCTTAGCGCAAGTTGCGAGCATGACAGAAATTTTAA	259	
QY	585	TCCGCTCAACTCCATGAAACGCGACCGCGCGCTCCAGCATTCACGACATCACAGCG	644	
Db	260	TCCGCTTAGCTCTGTAACAAAATAGAGGCGATTAAGCATTCATGATATCATCAACG	319	
QY	645	TCACGCGCGCGCATC	659	
Db	320	TAGCGATGAGAGACC	334	

```

RESULT 5
US-09-640-211A-1666
Sequence 1666, Application US/09640211A
Patent No. 6833446
GENERAL INFORMATION:
APPLICANT: Wood, Marion
APPLICANT: Shenk, Michael A.
APPLICANT: McGrath, Annette
APPLICANT: Glenn, Matthew
TITLE OF INVENTION: Compositions and Methods for the
Title of Invention: Modification of Gene Transcription
FILE REFERENCE: 11000.1021CIU
CURRENT APPLICATION NUMBER: US/09/640,211A
CURRENT FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 2368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1666
LENGTH: 422
TYPE: DNA
ORGANISM: Pinus radiata
US-09-640-211A-1666

```

	Query Match	7.9%	Score 105	DB 4	Length 422
	Best Local Similarity	68.9%	Pred. No. 1.3e-11		
	Matches 144	Conservative 0	Mismatches 65	Indels 0	Gaps 0
Qy	437	AAGGCATCCCATGACGAGAGACACAGCGTGTCTGCGGCTGACAAATTC	496		
Db	7	AAGGTTGTTCCTCGACTGAAGAAAGACACAGCGATTTTGTATGCGCTTCGCAAGTAC	66		
Qy	497	GGCAAGGGGGGACTGGCGGAGCATCTCGGCACTTCGTCATCTCCGGAGCGCAACGCGAG	556		
Db	67	GGCAAGGCGACTGGAGAAAGTATTTCTTMAAACTTTGTGTGTCAAGACACCAACCAA	126		
Qy	557	GTCGGAGCCACGCGCAGAAAGTACTTCATCCGCTCAATCCATGAACCGCAGCGCGCG	616		
Db	127	GTTGCGACCATGCTCAAAAGTACTACATTCGGCTTGCGATTAATMAAAACAAGAGA	186		
Qy	617	CGCTTCAGCATTCACGATCATCCAGCGT	645		
Db	187	AGATTCAGCATATGATATATCACTGT	215		

RESULT 6  
US-09-640-211A-2058

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	7.94	Score 105	DB 4	Length 436
Best Local Similarity	68.9%	Pred. No. 1	3e-11	
Matches 144	Conservative 0	Mismatches 65	Indels 0	Gaps 0

  

QY	437	AAGGCATCCCATGACGAGAGACACAGCTGTTCTTGCTGGGCTGCAAGTTG	496
Db	5	AAGGTTGTTCCCTGGACTGAAGAGACACAGCTGTTTGATGGGCTTCGCAAGTAC	64
QY	497	GGCAAGGGGGACCTGGCGGAGCATCTCGGCAACTCGTCATCTCGGGAGCGCAACGAC	556
Db	65	GGCAAGCGCAGCTGGAGAGATTTTCTAGAACTTTGTTGTCAAGACACCAACCAA	124
QY	557	GTGGCAGCCACGGCGCAGAACTACTTCATCCGCTCAATCCGATGAACCGCAGCCGCGC	616
Db	125	GTGGCAGCCAGTGTCAAAAGTACTACATTCGGCTTGTCGATATATAAAAAAGAGA	184
QY	617	CGCTCCAGCATCCACGACATCAACCGCT	645
Db	185	AGATCCAGCATACATGATATACCACTGT	213

```

RESULT 7
US-09-640-211A-1250
: Sequence 1250, Application US/09640211A
: Patent No. 683446
: GENERAL INFORMATION:
: APPLICANT: Wood, Marion
: APPLICANT: Shenk, Michael A.
: APPLICANT: McGrath, Annette
: APPLICANT: Glenn, Matthew
: TITLE OF INVENTION: Modifications and Methods for the
: TITLE OF INVENTION: Modifications of Gene Transcription
: FILE REFERENCE: 11000.1021C1U
: CURRENT APPLICATION NUMBER: US/09/640.211A
: NUMBER OF SEQ ID NOS: 2368
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1250
: LENGTH: 632
: TYPE: DNA
: ORGANISM: Eucalyptus grandis
US-09-640-211A-1250

```

Query Match	7.9%	Score 104.8	DB 4	Length 632
Best Local Similarity	52.5%	Pred. No. 1.5e-11		
Matches	275	Conservative	0	Mismatches 247; Indels 2; Gaps 2
Oy	83	GACGACAAGCGCTTCGAGAACGCGCTCGCGGCTTCGCGGCGCCCGCCGCGGACGGA	142	
Db	39	GCGCTCGCGCGCGTGGAACGCGCGGACGCCACGCGGCGACGACGCCGCGCTCTC	98	
Oy	143	GCGCGGCCGACGACGACTGCTTCGCGCGCTCGCGCGGAGCGTCCCGGGGCGAGGTTCG	202	





;; TITLE OF INVENTION: Compositions and Methods for the  
;; FILE OF INVENTION: Modification of Gene Transcription  
;; FILE REFERENCE: 11000.1021CIU  
;; CURRENT APPLICATION NUMBER: US/09/640,211A  
;; CURRENT FILING DATE: 2000-08-16  
;; NUMBER OF SEQ ID NOS: 2368  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 2038  
;; LENGTH: 1055  
;; TYPE: DNA  
;; ORGANISM: Pinus radiata  
US-09-640-211A-2038

Query Match 7.0%; Score 92.8; DB 4; Length 1055;  
Best Local Similarity 63.4%; Pred. No. 3.8e-09;  
Matches 142; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 413 TCCAGAGCGGAGCAGGAGGCGGCAAGGCGATCCCATGACGAGGAGAAAGCAGAGCTG 472  
DB 608 TCCAGCAATGACCGGAGAGGAAAAAGCGCTCCCATGCTCGAGGAAAGCAGAGATG 667  
QY 473 TTCTTGCTGGGGCTGAGCAAGTTGCGCAAGGGGAGCTGGCGAGCATCTCGCGCACTTC 532  
DB 668 TTCTTGATGAGCTGGAAGAAAGCTTGGAAGGGTACTGAGAGGATATCCAGAAATTTT 727  
QY 533 GTCATCTCGGCGAGCGCAAGCGAGGTGGCGAGCCGCGAGAAAGTACTTCATCCGCTC 592  
DB 728 GTGACGACGAGAAACACCCACAGTAGCCAGCATGCCAGAGTATTTTCTAAGGAG 787  
QY 593 AACTCCATGAACCGGACCGCGCGCTCCAGCATCCACGACAT 636  
DB 788 AGCAATCTTAATMAAGAAACGTGATCCAGTCTTTGATAT 831

RESULT 13  
US-09-640-211A-1369  
; Sequence 1369, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: 11000.1021CIU  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1369  
; LENGTH: 328  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-09-640-211A-1369

Query Match 6.9%; Score 92; DB 4; Length 328;  
Best Local Similarity 64.6%; Pred. No. 4.2e-09;  
Matches 137; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 425 CAGAGAGCGGAGGCGATCCATGAGCAGGAGAAAGCAGAGCTGTTCTGCTGGG 484  
DB 115 CAAGAGAAAGAAAGAGGATTCATGAGCAGAGGAGAAACAGAACTTTGTTGGGG 174  
QY 485 CTGAGCAAGTTCCGCAAGGGGAGCTGGCGAGCATCTCGCCACTTCGTCATCTCCGG 544  
DB 175 CTGAGAGGCTGGAGAGGGGATTTGAGAGGCACTCTAGAGACTATGAGCAAGA 234  
QY 545 ACGGCAAGCAGTGGCAGGCGCAGAGTACTTCATCCGCTCAACTCCATGAG 604  
DB 235 ACACCGGCGGAGGTGAGAGTCAATGCTCAAGAAATTTTCTCCGCGAAGTGAAGTTCAAC 294  
QY 605 CGGACGCGCGCGCTCCAGCATCCACGACAT 636

DB 295 AAGAAAAAGCGGCTCGAGCCTTTTGACAT 326

RESULT 14  
US-09-640-211A-1966  
; Sequence 1966, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: 11000.1021CIU  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1966  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-09-640-211A-1966

Query Match 6.9%; Score 91.4; DB 4; Length 427;  
Best Local Similarity 61.6%; Pred. No. 5.9e-09;  
Matches 146; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 400 CGGCAAGAGCTGCTCCAGGCGGAGCAGAGAGGCGGCAAGGCGATCCCATGACGAGCA 459  
DB 3 CGGCTCCCGCTCTCCATGCGCGCGTGAAGAGAAAGGTATACCATGACGAGCA 62  
QY 460 AGAGCAGGCTGTTCTGCTGGGGCTGAGCAAGTTGCGCAAGGGGAGCTGGCGAGCAT 519  
DB 63 AGAGCATCGAAGGTTTAAATTTGTTCTCCAGAAATTTGGTAAAGAGACTGGCGAGGAT 122  
QY 520 CTCGCGCACTTGCTCATCTCGCGAGCGCCAGCAGGTGGCGAGCCGCGAGAAATA 579  
DB 123 AGCTCGGACTTTGTGACTGTAAGAGACTCTCAAGTGGCAAGCATGCCAGAAATA 182  
QY 580 CTTGATCCGCTCAATCCATGAACCGGACCGCGCGCTCCAGCATCCAGCAT 636  
DB 183 TTATATCCGAGAGTAATGCTGCGGAGAAAGAGCGCTCCAGCCTTTTGACAT 239

RESULT 15  
US-09-640-211A-553  
; Sequence 553, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: 11000.1021CIU  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 553  
; LENGTH: 883  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-640-211A-553

Query Match 6.8%; Score 90.4; DB 4; Length 883;  
Best Local Similarity 63.2%; Pred. No. 1.1e-08;  
Matches 139; Conservative 0; Mismatches 81; Indels 0; Gaps 0;



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 17:21:58 ; Search time 5919 Seconds

(without alignments)  
10887.893 Million cell updates/sec

Title: US-10-630-636-1

Perfect score: 1330  
Sequence: 1 gtgcgagatccaccaccgca.....gacacaaaaaaaaaaaaa 1330

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_da:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_se:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1330	100.0	1330	AY151042	Oryza sat
2	1181.4	88.8	142680	AP004611	Oryza sat
3	1167.2	87.8	1329	AK068565	Oryza sat
4	851.4	64.0	53000	AP006233	Oryza sat
5	287.8	21.6	187154	OSJN00107	Oryza sat
6	198.4	14.9	1460	AK068623	Oryza sat
7	198.4	14.9	1465	AK106166	Oryza sat
8	197	14.8	1988	AK068138	Oryza sat
9	179	13.5	1502	AK111988	Oryza sat
10	171.6	12.9	1795	AK067964	Oryza sat
11	140.8	10.6	198102	AC145389	Zea mays
12	138.6	10.4	126532	AC130602	Oryza sat
13	135.6	10.4	138467	AC130612	Oryza sat
14	135.4	10.2	945	AY1519528	Arabidops
15	135.4	10.2	1161	AY086906	Arabidops
16	129.2	9.7	897	AY1519532	Arabidops
17	129.2	9.7	928	AY096571	Arabidops
18	128.2	9.7	1593	AY072090	Arabidops
19	128.2	9.6	1030	AK111571	Oryza sat

C	20	125.8	9.5	134982	8	AP003279	Oryza sat
	21	122.6	9.2	1115	8	HV0303354	Hordeum v
	22	120.6	9.1	1104	8	AK068595	Oryza sat
	23	120.6	9.1	1128	8	AK059813	Oryza sat
	24	119.2	9.0	1162	8	AK061500	Oryza sat
	25	119.2	9.0	1490	8	AK103432	Oryza sat
	26	118.6	8.9	867	8	AY1519533	Arabidops
	27	118.6	8.9	898	8	BT005473	Arabidops
	28	118.6	8.9	1640	8	AK118891	Arabidops
	29	116.4	8.8	897	8	AY1519529	Arabidops
	30	116.4	8.8	928	8	BT006122	Arabidops
	31	116.4	8.8	1645	8	AK118135	Arabidops
	32	115.8	8.7	909	8	AY196776	Malus xia
	33	115.2	8.7	63898	2	OSJN00045	AL732350 Oryza sat
C	34	115.2	8.6	150206	8	OSJN00083	AL606668 Oryza sat
	35	114.4	8.6	1204	8	AF239956	Hevea bra
	36	113.8	8.6	808	8	LR5243339	Lycopersi
	37	112.8	8.5	1208	8	HV0303355	Hordeum v
	38	110.6	8.3	1287	8	AY151044	Oryza sat
	39	110.6	8.3	1357	8	AK101062	Oryza sat
	40	110.2	8.3	1191	8	AB083027	Glycine m
C	41	110	8.3	100800	8	AC105260	Oryza sat
	42	109.2	8.2	142854	8	AC136216	Oryza sat
	43	109.2	8.2	1009	6	AX927169	Sequence
	44	109	8.2	1149	8	AK061427	Oryza sat
	45	109	8.2	3358	8	AK065594	Oryza sat

## ALIGNMENTS

RESULT 1  
AY151042  
LOCUS  
DEFINITION Oryza sativa (japonica cultivar-group) transcritpion factor MYBS1  
ACCESSION AY151042  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE  
1 (bases 1 to 1330)  
Lu C.A., Ho,T.H., Ho,S.L. and Yu,S.M.  
Three Novel MYB Proteins with One DNA Binding Repeat Mediate Sugar  
and Hormone Regulation of alpha-Amylase Gene Expression  
Plant Cell 14 (8), 1963-1980 (2002)  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
TITLE  
JOURNAL  
AUTHORS  
Lu,C.A., Ho,T.H., Ho,S.L. and Yu,S.M.  
Direct Submission  
Submitted (17-SEP-2002) Institute of Molecular Biology, Academia  
Sinica, Taipei, Nankang 115, Taiwan  
FEATURES  
source  
CDS  
1..1330  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/db\_xref="taxon:39947"  
20..940  
/note="OSMYBS1"  
/codon\_start=1  
/product="transcription factor MYBS1"  
/protein\_id="AA063152.1"  
/db\_xref="GI:24850303"  
/translation="MTGQAATTTTAAATAATREDKAFENALAACAPPADGAP  
DDWPAALASVPARSABEVRBVAIVDAIDGRVLPYRAGESSAAPDGAG  
AAASKDGRRPDERGCGGCGGCGSCAABERRKGIPTWEEHRLPLIGDKPG  
KGDNRSTSRNFVISRTPTOVASHOKTFRILNSNRNRKRSSIHDTITSVTRAGDOVAAO  
QGAPITGHQATGNPAALALGPPGKHHHHHPGAGAPPMPVSAAPMGHPVAGHMVPA  
AVGTPVVFPPGHAPVVFVVGYPAPPAKHQ"

ORIGIN	Query Match	Similarity	Score	DB #	Length	Matches	Conservative	0	Mismatches	0	Indels	0	Gaps	0
	Best Local	100.0%;	100.0%;		1,4e-131									
	Matches	1330;												
Qy	1	GTGGAATCCACCA	CGGATGACCTCC	CAGCGGCGCA	CGACGACGAC	CGGCGGCGGC	60							
Db	1	GTGGAATCCACCA	CGGATGACCTCC	CAGCGGCGCA	CGACGACGAC	CGGCGGCGGC	60							
Qy	61	GGCGGCGGCGT	GACGAGGAGGAG	CGACGACGAC	CGGCGGCGGC	CGGCGGCGGC	120							
Db	61	GGCGGCGGCGT	GACGAGGAGGAG	CGACGACGAC	CGGCGGCGGC	CGGCGGCGGC	120							
Qy	121	GGCGGCGGCGG	CGGCGGAGGCG	CGCGCGACG	CGACTGCTT	CGCGCGCTCG	180							
Db	121	GGCGGCGGCGG	CGGCGGAGGCG	CGCGCGACG	CGACTGCTT	CGCGCGCTCG	180							
Qy	181	GAGCGTCCCGG	GCGGAGAGTGG	CGGAGAGAGT	GGGCACTAC	GAGGCGCTGG	240							
Db	181	GAGCGTCCCGG	GCGGAGAGTGG	CGGAGAGAGT	GGGCACTAC	GAGGCGCTGG	240							
Qy	241	GAGCGTCCCGG	CGGCGGAGGCG	CGGCGGCGCT	CGCGGCTT	CGCGGCGGAG	300							
Db	241	GAGCGTCCCGG	CGGCGGAGGCG	CGGCGGCGCT	CGCGGCTT	CGCGGCGGAG	300							
Qy	301	CGCGGCGGCGG	CGGCGGAGGCG	CGGCGGCGCT	CGCGGCTT	CGCGGCGGAG	360							
Db	301	CGCGGCGGCGG	CGGCGGAGGCG	CGGCGGCGCT	CGCGGCTT	CGCGGCGGAG	360							
Qy	361	GCGGCGGCGG	CGGCGGAGGCG	CGGCGGCGCT	CGCGGCTT	CGCGGCGGAG	420							
Db	361	GCGGCGGCGG	CGGCGGAGGCG	CGGCGGCGCT	CGCGGCTT	CGCGGCGGAG	420							
Qy	421	GGAGAGGAGG	GCGGAGGAGT	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	480							
Db	421	GGAGAGGAGG	GCGGAGGAGT	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	480							
Qy	481	GGGCGTGGCA	AGTTCCGCAAG	GCGGAGGAGT	CGGAGGAGG	AGGAGGAGG	540							
Db	481	GGGCGTGGCA	AGTTCCGCAAG	GCGGAGGAGT	CGGAGGAGG	AGGAGGAGG	540							
Qy	541	GCGGAGGAGG	CGGAGGAGT	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	600							
Db	541	GCGGAGGAGG	CGGAGGAGT	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	600							
Qy	601	GAAAGGAGG	CGGCGGCGCT	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	660							
Db	601	GAAAGGAGG	CGGCGGCGCT	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	660							
Qy	661	GATGCGCGG	CGGAGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	720							
Db	661	GATGCGCGG	CGGAGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	720							
Qy	721	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	780							
Db	721	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	780							
Qy	781	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	840							
Db	781	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	840							
Qy	841	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	900							
Db	841	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	900							
Qy	901	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	960							
Db	901	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	960							
Qy	961	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	1020							
Db	961	GCGGCGGCGG	CGGCGGAGG	CGGAGGAGG	AGGAGGAGG	AGGAGGAGG	1020							

Qy	1021	CGGCTCATCGT	CCCTGATCATCTT	GTCTCTCTCA	CAATCTTCA	CTCAATGTA	TAACATACATA	1080
Db	1021	CGGCTCATCGT	CCCTGATCATCTT	GTCTCTCTCA	CAATCTTCA	CTCAATGTA	TAACATACATA	1080
Qy	1081	GATCAGATG	CCCAAGAGG	CGAGGATTT	GGGATTTAA	AGCCAAAT	TAAGTAATTTTGGT	1140
Db	1081	GATCAGATG	CCCAAGAGG	CGAGGATTT	GGGATTTAA	AGCCAAAT	TAAGTAATTTTGGT	1140
Qy	1141	GACTGTTT	GCAAGATG	ATCATACG	TACACCCG	GGTGAAGT	TAGCTCCAAATGTGATGT	1200
Db	1141	GACTGTTT	GCAAGATG	ATCATACG	TACACCCG	GGTGAAGT	TAGCTCCAAATGTGATGT	1200
Qy	1201	AATTAGCAG	CGGCGCTT	CCGTAAGT	GTGGCGG	CGGATGAT	GATCTTCCAGGGGTTGCAAT	1260
Db	1201	AATTAGCAG	CGGCGCTT	CCGTAAGT	GTGGCGG	CGGATGAT	GATCTTCCAGGGGTTGCAAT	1260
Qy	1261	TAGGATTT	GATTTTCC	ATTTTGTG	ATGTAATTT	GGCAAC	TGTTCTCATTTGGACCAAAAA	1320
Db	1261	TAGGATTT	GATTTTCC	ATTTTGTG	ATGTAATTT	GGCAAC	TGTTCTCATTTGGACCAAAAA	1320
Qy	1321	AAAAAAAA	1330					
Db	1321	AAAAAAAA	1330					
RESULT 2	AP004611	142680 bp	DNA	linear	PLN 31-AUG-2004			
LOCUS	AP004611	Oryza sativa (japonica cultivar-group)	genomic DNA, chromosome 1,					
DEFINITION	BAC clone: O01005_B10.							
ACCESSION	AP004611	BA000010						
VERSION	AP004611.3	GI:21902089						
KEYWORDS								
SOURCE	Oryza sativa (japonica cultivar-group)							
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.							
REFERENCE	1	Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Nishimura, Y., Cheng, Z., Negamatsu, Y., Antonio, B. A., Kanamori, H., Hosokawa, S., Masukawa, M., Arita, K., Chuden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Hijiishi, S., Honda, M., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikono, M., Ito, S., Ito, Y., Ito, Y., Iwabuchi, A., Kamiya, K., Katsura, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizudayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Teuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yokawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J. H., Kim, H. I., Eun, M. Y., Yano, M., Jiang, J. and Gojobori, T.						
TITLE	The genome sequence and structure of rice chromosome 1							
JOURNAL	Nature 420 (6913), 312-316 (2002)							
MEDLINE	22337376							
PUBMED	12447438							
REFERENCE	2 (bases 1 to 142680)							
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.							
TITLE	Direct Submission							
JOURNAL	Submitted (10-JUN-2002) Takuji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan							
COMMENT	(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rpg.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)							
	On Jul 17, 2002 this sequence version replaced gi:1973519.							
	Genes were predicted from the integrated results of the following: GENSCAN 1.0, BLASTN 2.0, BLASTX 2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant Protein database, nr							
	(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at							



```

CDS
      /gene="OJ1005_B10.11"
      /complement(join(39042..39137,40177..40346,40543..40590,
      41257..41417,42100..42248))
      /gene="OJ1005_B10.11"
      /note="hypothetical protein"
      /protein_id="BAC05646.1"
      /db_xref="GI:2190209"
      /translation="MARTGAVAHARTDGLATFNPMTKKYLEGYCTSDPEPRAH

Query Match      88.8%; Score 1181.4; DB 8; Length 142680;
Best Local Similarity 91.2%; Pred. No. 2e-116;
Matches 1316; Conservative 0; Mismatches 6; Indels 121; Gaps 1;

Qy      1 GTGCGAGATCCACCAACCCGATGACCTCCAGGCGGCGACGACGACGACGCGCGGCGC 60
Db      138939 GTGCGAGATCCACCAACCCGATGACCTCCAGGCGGCGACGACGACGCGCGGCGC 138998

Qy      61 GCGCGCGCGCTTGACACAGGAGAGACGACAAAGCGCTTGAGAAACGCGCTCGCGGCTTGCGC 120
Db      138999 GCGCGCGCGCTTGACACAGGAGAGACGACAAAGCGCTTGAGAAACGCGCTCGCGGCTTGCGC 139058

Qy      121 GCGCGCGCGCGCGCGGACGAGAGCGCGCGCGACGACGACGACGACGACGACGACGACGACG 180
Db      139059 GCGCGCGCGCGCGCGGACGAGAGCGCGCGCGACGACGACGACGACGACGACGACGACGACG 139118

Qy      181 GAGCGTGCCCGGAGGCGAGAGTCCGCGGAGAGAGTCCGAGGACCTACGAGGCGCTGTGGA 240
Db      139119 GAGCGTGCCCGGAGGCGAGAGTCCGCGGAGAGAGTCCGAGGACCTACGAGGCGCTGTGGA 139178

Qy      241 GGAGCTCGCGGCGCATCCACGCGGCGCGCGCTCCCGGCTACGCGCGGCGGAGAGATC 300
Db      139179 GGAGCTCGCGGCGCATCCACGCGGCGCGCGCTCCCGGCTACGCGCGGCGGAGAGATC 139238

Qy      301 GCGCGCGCGCGCGCGGACGAGCGCGCGCGCGCGCGCGCTCCAAAGACGCGGACACCG 360
Db      139239 GCGCGCGCGCGCGCGGACGAGCGCGCGCGCGCGCGCGCTCCAAAGACGCGGACACCG 139298

Qy      361 GCGCGACGAGCGCAAGGCGCGCGCGCGCGGATCGACGCGCGCAAGAGTCTCCAAAGC 420
Db      139299 GCGCGACGAGCGCAAGGCGCGCGCGCGCGGATCGACGCGCGCAAGAGTCTCCAAAGC 139358

Qy      421 GGAGCAGAGAGAGGCGCAAGGGGCTATCCATGACGAGAGAAAGCA----- 465
Db      139359 GGAGCAGAGAGAGGCGCAAGGGGCTATCCATGACGAGAGAAAGCAAGGTAACGCGCATT 139418

Qy      466 ----- 465
Db      139419 CCGGAATCAAGACGAGACGCTGTTGTTCTTCTTCTCGGAACAGCTCGATTGC 139478

Qy      466 ----- 465
Db      139479 GAATTCATGCGCGGACGCTGACTCGATTGCGATTGATTGCGCGCAGGCTGTTCTTGC 139538

Qy      480 TGGGGCTTGACAAATTCCGGAAGGGGGAATGCGGAGCATCTCGGCAACTTCGTCATCT 539
Db      139539 TGGGGCTTGACAAATTCCGGAAGGGGGAATGCGGAGCATCTCGGCAACTTCGTCATCT 139598

Qy      540 GCGGACGCAACGAGGTGCGAGCGACGCGCGCAAGTAATTATTCGCTCAACTCA 599
Db      139599 GCGGACGCAACGAGGTGCGAGCGCGCGCGCAAGTAATTATTCGCTCAACTCA 139658

Qy      600 TGAACCGGACCGCGCGCTCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATC 659
Db      139659 TGAACCGGACCGCGCGCTCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATC 139718

Qy      660 AGGTGCGCGCGCAGAGAGGCGCGCGGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATC 719
Db      139719 AGGTGCGCGCGCAGAGAGGCGCGCGGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATC 139778

Qy      720 GCGCGCGCGCTGAGGCGCGCGGAGCATGAAGCAACACCAACCAACCAACCGCGGCGCGCGC 779
Db      139779 GCGCGCGCGCTGAGGCGCGCGGAGCATGAAGCAACACCAACCAACCAACCGCGGCGCGCGC 139838
```

```

Qy      780 GCGCGCCCATGCCCATTATACAGCGCGCGCGCCCATGCGCCACCCCGTGGCGGCCACATGG 839
Db      139839 GCGCGCCCATGCCCATTATACAGCGCGCGCGCCCATGCGCCACCCCGTGGCGGCCACATGG 139898

Qy      840 TGGCCCGCGCGCGTGGGACGCGCGGTGGTGTCCCGCGCGGACGACGCGCGTACGTCGTGC 899
Db      139899 TGGCCCGCGCGCGTGGGACGCGCGGTGGTGTCCCGCGCGGACGACGCGCGTACGTCGTGC 139958

Qy      900 CCGTGCGCTACCGCGCGCTCCGCGCAAGATGACCAATGACGCGCGCATGAGCGGACATG 959
Db      139959 CCGTGCGCTACCGCGCGCTCCGCGCAAGATGACCAATGACGCGCGCATGAGCGGACATG 140018

Qy      960 AGCAGCATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
Db      140019 AGCAGCATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 140078

Qy      1020 CCGGCTCATGTCCTCGATCATCTGTTCTTCTCAAAATCTCAATGTAATGTAATCAT 1079
Db      140079 CCGGCTCATGTCCTCGATCATCTGTTCTTCTCAAAATCTCAATGTAATGTAATCAT 140138

Qy      1080 AGATCAGATCCAAAGATGACGAGGATGGGATTTAAAGCGAATTAAGTAATTTTGC 1139
Db      140139 AGATCAGATCCAAAGATGACGAGGATGGGATTTAAAGCGAATTAAGTAATTTTGC 140198

Qy      1140 TGACTGTTTGCAAGTATCATACGTAACCCGCTGAACCTTAAGCTTCCAAATGTGATG 1199
Db      140199 TGACTGTTTGCAAGTATCATACGTAACCCGCTGAACCTTAAGCTTCCAAATGTGATG 140258

Qy      1200 TAATTACAGCGGCGCTCCGATGCGTGGCGCGCATGATGATCTTGACGAGGTTTCA 1259
Db      140259 TAATTACAGCGGCGCTCCGATGCGTGGCGCGCATGATGATCTTGACGAGGTTTCA 140318

Qy      1260 TTAGGATTTGATTTTCCATTTTGTGATGATTAATTTGCAACGTCTCATTTGACCAAAA 1319
Db      140319 TTAGGATTTGATTTTCCATTTTGTGATGATTAATTTGCTACTGCTCATTTGACCAAAAAC 140378

Qy      1320 AAA 1322
Db      140379 ATA 140381

RESULT 3
AK068565
LOCUS
DEFINITION
ACCESSION
AK068565
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 The Rice Full-length cDNA Consortium, National Institute of
  Agrobiological Sciences Rice Full-length cDNA Project Team,
  Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
  Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
  Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
  Onshiki,K., Shishiki,T., Foundation of Advancement of International
  Science Genome Sequencing & Analysis Group, Otsu,Y., Murakami,K.,
  Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
  Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
  Marikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nishikura,J.,
  Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
  Kusumegi,T., Oka,M., Ryu,K., Ueda,M., Matsubara,K., RIKEN,
  Kawai,J., Carninci,P., Adachi,U., Aizawa,K., Arikawa,T., Fukuda,S.,
  Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
  Kagawa,I., Kondo,S., Kono,H., Miyazaki,A., Otsu,N., Oka,Y.,
  Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
  Yoshino,M., and Hayashizaki,Y.
```



QY 1081 GATCAGATGCCAGAGTGCAGGGATTGGGATTAAAGCCGATTAAGTATTTGGCT 1140  
DB 1096 GATCAGATGCCAGAGTGCAGGGATTGGGATTAAAGCCGATTAAGTATTTGGCT 1155  
QY 1141 GACTGTTGGAAGGATCATCATACCCCGGGAAGGCTTACTCCAAATGTTGATGT 1200  
DB 1156 GACTGTTGGAAGGATCATCATACCCCGGGAAGGCTTACTCCAAATGTTGATGT 1215  
QY 1201 AATTAGCAGGCGCTTCGTACGTGTGCGCCGATGATGATCTTCAGGGGTTGCAAT 1260  
DB 1216 AATTAGCAGGCGCTTCGTACGTGTGCGCCGATGATGATCTTCAGGGGTTGCAAT 1275  
QY 1261 TAGGATGATGATTTCCATTTTGGTATTAATTTGCCAATCTGTCTCATTTGGACC 1314  
DB 1276 TAGGATGATGATTTCCATTTTGGTATTAATTTGCCAATCTGTCTCATTTGGACC 1329

RESULT 4  
AP006233 53000 bp DNA linear PLN 31-AUG-2004  
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,  
DEFINITION BAC clone: B1249E06, complete sequence.  
ACCESSION AP006233 BA000010  
VERSION AP006233.2 GI:29824107  
KEYWORDS HTG.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehharitoidae; Oryzeae; Oryza.

REFERENCE  
AUTHORS 1  
TITLE Sasaki, T., Matsumoto, T. and Katayose, Y.  
JOURNAL Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
AUTHORS clone: B1249E06  
TITLE Published Only in Database (2003)  
2 (bases 1 to 53000)  
DIRECT SUBMISSION  
SUBMITTED (19-FEB-2003) Takuji Sasaki, National Institute of  
Agricultural Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: sasaki@nias.affrc.go.jp; URL: http://rgp.dna.affrc.go.jp/  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
On Apr 11, 2003 this sequence version replaced gi:28460672.  
The orientation of the sequence is from -21M13 to M13rev of the BAC  
clone.

FEATURES  
source location/Qualifiers  
1..53000  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="1"  
/clone="B1249E06"

ORIGIN  
Query Match 64.0%; Score 851.4; DB 8; Length 53000;  
Best Local Similarity 98.7%; Pred. No. 1.7e-81;  
Matches 858; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

DB 230 CATCACCAGGTCACCGCCGCGATCATGATGTCGCGGACAGAGGCGCCCGATCACCGG 289  
QY 694 CCACAGGCGCAGGCGAACCCTCGGCGGCGGCTGAGGCGCCCGGAGCATGAAGACCA 753  
DB 290 CACAGGCGCAGGCGAACCCTCGGCGGCGGCTGAGGCGCCCGGAGCATGAAGACCA 349  
QY 754 CCAACACACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 813  
DB 350 CCAACACACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 409  
QY 814 GGGCGACCCCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 873  
DB 410 GGGCGACCCCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 469  
QY 874 GCGGCGCAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 933  
DB 470 GCGGCGCAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 529  
QY 934 CCAATGACGCGCGCATGAGCAGACATGAGCAGCATTTCTTCCTCTCTCTCTCTCTCT 993  
DB 530 CCAATGACGCGCGCATGAGCAGACATGAGCAGCATTTCTTCCTCTCTCTCTCTCTCT 589  
QY 994 ATCTGATTTGTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1053  
DB 590 ATCTGATTTGTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 649  
QY 1054 ACAATGCACTTAATGTAATCATATGATGATGATGATGATGATGATGATGATGATGAT 1113  
DB 650 ACAATGCACTTAATGTAATCATATGATGATGATGATGATGATGATGATGATGATGAT 709  
QY 1114 AAAAGCGAATTAATGTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1173  
DB 710 AAAAGCGAATTAATGTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 769  
QY 1174 TGAAGCTTGTCTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1233  
DB 770 TGAAGCTTGTCTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 829  
QY 1234 GATCGATGATCTTGCAGGCGGTTGCAATTAAGGATTTGTTGTTGTTGTTGTTGTTGTT 1293  
DB 830 GATCGATGATCTTGCAGGCGGTTGCAATTAAGGATTTGTTGTTGTTGTTGTTGTTGTT 889  
QY 1294 TGCCCACTGTTCTATTTGACCAAAAAA 1322  
DB 890 TGCCCACTGTTCTATTTGACCAAAAAA 918

RESULT 5  
OSJN00107/c 187154 bp DNA linear PLN 10-FEB-2004  
LOCUS Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNB0062B06,  
DEFINITION complete sequence.  
ACCESSION AL606729 GI:32488976  
VERSION AL606729.3  
KEYWORDS HTG.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehharitoidae; Oryzeae; Oryza.

REFERENCE  
AUTHORS 1  
PENG, Q., ZHANG, Y., HAO, P., WANG, S., FU, G., HUANG, Y., LI, Y., ZHU, J.,  
LIU, Y., HU, X., JIA, P., ZHANG, Y., ZHAO, Q., YING, K., YU, S., TANG, Y.,  
WANG, Q., ZHANG, L., LU, Y., MU, J., LU, Y., ZHANG, L. S., YU, Z., FAN, D.,  
LIU, X., LU, T., LI, C., MU, Y., SUN, T., LEI, H., LI, T., HU, H., GUAN, J.,  
WU, M., ZHANG, R., ZHOU, B., CHEN, Z., CHEN, Z., JIN, Z., WANG, R.,  
YIN, H., CAI, Z., REN, S., LV, G., GU, W., ZHU, G., YU, Y., JIA, J.,  
ZHANG, Y., CHEN, J., KANG, H., CHEN, X., SHAO, C., SUN, Y., HU, Q.,  
ZHANG, X., ZHANG, W., CHEN, L., CHENG, C., SHENG, H., GU, J., CHEN, S.,  
MI, L., ZHU, F., CHEN, W., LAN, L., LAI, Y., CHENG, Z., GU, M., JIANG, J.,  
LI, J., HONG, G., XUE, Y. and HAN, B.  
Sequence and analysis of rice chromosome 4

TITLE

JOURNAL Nature 420 (6913), 316-320 (2002)  
MEDLINE 22373737  
PUBMED 12447439  
REFERENCES  
AUTHORS  
2  
Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X.,  
Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J., Zhang, L.,  
Liu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C.,  
Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K., Zhou, B.,  
Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P., Fu, G.,  
Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F.,  
Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y.,  
Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J.,  
Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and  
Hong, G.F.  
Title Direct Submision  
Journal Submitted (08-SEP-2001) Han Bin, National Center for Gene Research,  
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,  
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:  
bhan@ncgr.ac.cn  
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC  
clone: OSJNB0062B06.  
On Jul 9, 2003 this sequence version replaced gi:21912495.  
Web site: http://www.ncgr.ac.cn  
----- Summary Statistics  
Assembly program: phrap  
-----  
This is a complete sequence.  
Genes were identified by a combination of several methods: Gene  
prediction programs including Fgenesh (http://www.softberry.com/),  
GENSCAN (http://CCR-081.mil.edu/GENSCAN.html), GenemarkHMM  
(http://genemark.biology.gatech.edu/genemark/), tRNAscan-SE (Sean  
Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the  
complete sequence against NCBI non-redundant protein database (nr)  
(ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.  
Location/Qualifiers  
1. 187154  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/sub\_species="japonica"  
/db\_xref="taxon:39947"  
/chromosome="4"  
/clone="OSJNB0062B06"  
/clone\_1fb="CUGI-OSJNB"  
3932..5799  
/gene="OSJNB0062B06.1"  
join(3932..4688,5168..5282,5395..5761,5847..6161,  
6241..6561,6647..6799)  
/gene="OSJNB0062B06.1"  
/codon\_start=1  
/protein\_id="CAE04043.2"  
/db\_xref="GI:38345960"  
/translation="MTNLAEVYVMVHGVGLPLVGIVEFLHGTCTYFRRFOAIL  
PSMKNLSIFGTFMOKLEBLRKKMKRRLAAVOGQOQRFILCODKAGRIYRKYK  
OEVLKANGTCHGSCAKPKLHRPCTHYIAAEEGIDPAVVSQKAIYHTWS  
EIVGIGLGEPTENDEVLTNIPDSKLGKGRAPRTIRNDMDSDGKRYKSCKD  
ERGHYKCPDKKEKPNRAEAGLSSAVDAGAPTDGTGTYTSVGPGRSPTPSPYC  
GTGSDWVPGTDVRLREAGLLPMCLVEAAGDVDPARWTVDSLLAALVDRRPE  
THTPLPCGEVAPLTLDISYLLGLADAVRPVTVGWDQDLFARFALVORAPHLP  
LEPLAHNRTGPTKRLMLOFTVEQLADAEYSYRCLAEVILMLFGVMCGGHGH  
YKSYRLSMARRGVEDPNMRSDCVYLPLPMCOMPLDGRCOLMASRNDIRGRH  
FKQYVNRRTTRTACATYEWVDTEPLDITTCLOBRKWFAASSTQYQORXAYE  
ROCRBOQSDMRVLTALPPEWAKPRCRGCDROVRSIKPTTLGRFFVCALITDDDF  
MEPPRCQYREMDITRRVLTTPSHVQLELPQYKVTARFERGSGSSRRG"  
gene  
CDS  
complement(join(10043..10336,10864..11112,13427..13562,  
16168..16386,17655..18033,18577..18817))  
/gene="OSJNB0062B06.2"  
/codon\_start=1  
/protein\_id="CAE04044.2"  
/db\_xref="GI:38345961"

/db\_xref="UniProt/TREMBL:Q7X7F4"  
/translation="MERADKLKQDVRILFETRNDAKMKMLVDVAQRLGIDHLPDEI  
NNVISDIKSEFTSSSLHEVALHFRLLREHINWSPVFEKFGKDGCFITIDEPRL  
ALLSLVNAHLIVDEPRLERAMFARMHDSGLKALPNOINRALPLPLPT  
YKRYEMLHMYEYQSECTIVYLLDLDALENNLQGHNLKELKAFNOMKWDLYVEL  
SHVRDRAVESTLWSYALFEENLTLTNLAKLIIIVFLVMDTDDHATIECRKLN  
AIOFMDBSASILPEYMKKPYRALQNFRETEAVQSDKRYVCMKPEFNLSYVL  
OEPEMLQNYSPAFKERYALSTSTSYPLCTVTAVGQDAVTESFELTVRSAAVI  
AKATIMEPMDIAFKSGRKNKGGAANTVECYNNENKTSVALDIEISMIESWRTL  
NQRCDHCOOPVQVRLNLNLAIVSPFFYDKKDAVTSRYIQEIVGDLFVNPVP1"  
26861..27649  
/gene="OSJNB0062B06.3"  
26861..27649  
/gene="OSJNB0062B06.3"  
/codon\_start=1  
/protein\_id="CAE04045.2"  
/db\_xref="GI:38345962"  
/translation="MVFDLNDLFLPRADLEVLTDMFSTQEISSLITLIPADKVPDQ  
FNGFPMKCMHIIIAHFAEATDLOVANNSTYTLTAVGQDAVTESFELTVRSAAVI  
LTKLMADRLQGVLEVVGENQYEFIKRTIOCLAQAEYIHOCQKREIILKLD  
EKADPTIEHTAISVMENMGPRKMLMWMQAFSSASTILLNGVSGISFRCXKGVQ  
GDPISPLFVGAELQRIYRARFNLGKLPFTYGLRVLQDQK"  
complement(36241..36855)  
/gene="OSJNB0062B06.4"  
complement(36241..36855)  
/gene="OSJNB0062B06.4"  
/codon\_start=1  
/protein\_id="CAE04046.2"  
/db\_xref="GI:38345963"  
/translation="MPTTVAORGEIAARRGRPEESSTRNGKDDGTCVGCARP  
RVTLPLRCRRRRRRPAHGGGCGSGCRDGGTGGGLRTNLRACGGRRR  
GAGGDDDVGRSGAATGRLAAGCMRGCRGRCGVHARLRKRENRGHHFI  
GAGRPATGRAGCGNGRSMKLGSEIBEGGIEPLINLE"  
complement(45605..50168)  
/gene="OSJNB0062B06.5"  
complement(45605..50168)  
/gene="OSJNB0062B06.5"  
complement(join(45605..45901,46925..47173,47477..47612,  
48055..48283,49052..49430,49925..50168))  
/gene="OSJNB0062B06.5"  
/codon\_start=1  
/protein\_id="CAE04047.2"  
/db\_xref="GI:38345964"  
/translation="MVRAEKLKQDVRILFETRNDAKMKMLVDVAQRLGIDHLPDEI  
NYSLSINASEFTSSSLHVDYALFRLRQGRFVSPVFEKFGKDDGCTGINTDP  
RGLSLVNAHLIHDDELEALSFATOLHSLSSGDDNLPHLIDQINRLDVLPLR  
TYRMETLCVMEYRQEGHILPILHLEMLFNLOHVLKELKAISEWMDLYGML  
LSYLRDRVESYVWSYVYFVEDSALARMITFKIIAFLIIMDDTYDSVATIOECKLN  
EAIORMDSATAPLPEYIKKYSALTKTFKEFEIHHVEDGQYRIDHTKARONLSAY  
LOEAEWSYQNKPSPEQVALSYTSYPLCVSTVGGALTNAPFAMANDIGAK  
IACKITRFNDIAPFKORGRKRGDVSTTECYNNENKTSVALDIEISMIESWRTL  
TINOLCEHRELLAFOOVNLALICATFFYGRKDAYTPSTHDETVESLFRVPVSI"  
complement(64764..71118)  
/gene="OSJNB0062B06.6"  
complement(join(64764..64805,64958..65940,66221..67084,  
67774..67858,68435..68536,68753..69149,69306..69546,  
69909..70953,71089..71118))  
/gene="OSJNB0062B06.6"  
/codon\_start=1  
/protein\_id="CAE04048.2"  
/db\_xref="GI:38345965"  
/translation="MTYLRVNPAPAPAPADVSNSSAVTTPKQGLSKTAGVSYA  
KRLOPVADSLPPEGPSKVSASSSVLKSRIIVSVHSPEDASVLSBSSWSSEKAK  
YMKRKHSSRSSTLSETRLRRAHFLVMEGRFCRLSESHRVIDPLXLSVSVG  
LDMKSSAPRRDLALNPOMLPLPLCFQOLISPHSLVQASLRLQASRRRPPMSS  
WCAGOTSPRLPLPLGRSSATMBDLICVPLRGMVLTTRREGOMRALARPLMTRFARLH  
RDOUWKYVPEDESVTLSPRADAHILICORLDCNCHGRKILMEVVAHILAIAPBPFGV  
YRARCMEGIPMAHRTAVAKLIGRSPVLCGGRKKILMEVVAHILAIAPBPFGV  
VSRAGRAVTVGSMIKASTTRTTVAATATKPLPEEAANTPGAELRKFLVPLQVR  
VOHSSSPSHRLKPTQATNLNINSPCNVGRPHSKSPRGAGALLLVPSPELFFLHQ  
DNCNGQHLDDPMLAEAVINYSPDGPEQGSPPVAFPAJBALAVLQGSVNASDLLA



SLATPRFKQPTPRTSLMADLYEKGCCCKLKIVANGANNFFLPADGTREGIILAA  
 DDNYRLRSQSEETNLSPEEALARSIAKATLGAALANRIKIRORSRLTWIKEDANT  
 KLPHIKANARRKNFIHYLOTSTGEFAISHOKEDELFNPSDRKQOSRELSMOS  
 LGMSFPLSDOEELITMEBELTITKSPSKAPPEPDSIFSAFYKCMDLIKHDLFNAY  
 RSFTKLNONOLDLNTCHICLIPKKADASAPEHRPISLIHSFKIISKLANLAR  
 MAELVTANQSAFIRKRAIHDFNLYVQULVKFRKRNKPTLMKHIDISADAYNAYM  
 MEVLVSJDAOESIFNFKOLNIFGASGLQNNVHTQULPLACBSLQVYIPQDIPRL  
 ADPFTSYLGLPLHTRERLKIIDFAPLIDKIGARTLSWKGEFSSGRTETLYKYLSTIP  
 TYHLSAIPIPKMAIKKIDIKIRAPLWKGEBENVKPGASLVMPVCRPTMGGLGIL  
 DLHKEFTTLRVMILRMWNSKPDQMEDESDILHFOACRISIGNHKSLFWE  
 DNMLQVCPKDVSPNLFKLAMKSRRTKNELNSMLSPROITNIEIHELVOIGM  
 LQHTHTSSMEDKLEMTQSSGITYTASATYLAQFQSGEPFNDNFMGLCGKANNAGSG  
 "

gene  
 complement (72467, .77023)  
 /gene="OSJNB0062B06.7"  
 /complement {join (72467, .72757, 72908, .73132, 73257, .73392,  
 74671, .74889, 75097, .75484, 76780, .77023)}  
 /gene="OSJNB0062B06.7"  
 /protein\_id="CAE04049.2"  
 /codon\_start=1  
 /protein\_id="CAE04049.2"  
 /db\_xref="GI:38345966"  
 /db\_xref="GI:38345966"  
 /translation="MRKRADELKEKRTOLGTCEDIVGTNNLVNVAIOHIGIEHFKOE  
 IDNTRIRTSSEFTSSSLHRYALMFLRLREHGLMSPDVFGKREFDDDLRLSVTID  
 HDTRGLSLVNAHLVHGEPELEALSIARHLKSMTRDCLPFLANQVRLNLTA  
 LPRCKRLSTSLFISEYOESEIILELAKLDFNVQVHLMELKSIBEMWDLTA  
 YVGLNVARDRABEGYLSCLFYEKDLSTFTEPAKMIILVTLDDPFDHATIOER  
 QUNSAIORMDESAITLPEYLKREYRELNRKFLQDOVDNDEKRYATYRKEKOS  
 TYVLOEAPRSGDQITLTJAMS SVPLICVSGTGMGVYDNTMFWVASRTYAVASAK  
 IGRPMNDIAAMKRKNGKDVASVECTMNEKHTMEVAIDIKDISLVDENRTLNQALF  
 EDHLKPFVEQVNLTAJMSASFYDERDAPFLPDLDTLESLEVNVP1"

Query Match 21.6%; Score 287.8; DB 8; Length 187154;  
 Best Local Similarity 76.3%; Pred. No. 3,4e-22;  
 Matches 439; Conservative 0; Mismatches 87; Indels 49; Gaps 5;

Qy 461 GAGCAAGGCTGTTCTGCTGGGCTGGACAAGTTCCGCAAGGGGACTGGCGAGCATC 520  
 Db 57943 GTGCGCAGGCTGTTCTGCTGGGCTGGACAAGTTCCGCAAGGAGACTTTCACAGATC 57884

Qy 521 TCGGGCAACTGTCATCTGCGGACGCAAGGAGGTGGAGCCACCGCGCAGAAGTAC 580  
 Db 57883 TCGGGCAACTGTCATCTGCGGACGCGGACGAGGAGGTGGAGCCACCGCAGAAGTAC 57824

Qy 581 TTCTCCGCTCACTCCATGAACCGGACCGCGCGCTCCAGCATCCAGACATCAC 640  
 Db 57823 TTCTCCGCTCACTCCATGAACCGGACCGCGCGCTCCAGCATCCAGACATCAC 57764

Qy 641 AGCGTCAACCGCGGATCAGGTGGCGGACGAGGAGGCGCGCGCATCCGCGCAC 700  
 Db 57763 AGCGTCAACCGCGGCG--ACGTTGCGCGGACGAGGAGGCGCGCGCATCCGCG 57717

Qy 701 GCGACGGGCAACCGCGGCGGCGCGGCGTGGCGCGCGCGGCGCATGAAGCACACAC 760  
 Db 57716 CCGACGGGCAACCGCGGCGGCGGCGGCGTGGCGCGCGCGGCGCATGAAGCAC 57665

Qy 761 CACCAACCGCGGCGGCGCGCGCGCATGTCATGTCAGCGCGCGCGCATGCGCAC 820  
 Db 57664 CACCAACCGCGGCGGCGCGCGCGCATGTCATGTCAGCGCGCGCGCATGCGCAC 57618

Qy 821 CCGTCCGCGGCGCATGTCGCGCGCGCGTGGCGACGCGCGGTGTGTTCCCGCGGCG 880  
 Db 57617 CCGTCCGCGGCGCATGTCGCGCGCGCGTGGCGACGCGCGGTGTGTTCCCGCGGCG 57560

Qy 881 CACCGCGCGTACGTGTCGCGCGCGTGGCGTACCGCGCGCGTCCGCGCAAGATCCACATGA 940  
 Db 57559 CACCGCGCGTACGTGTCGCGCGCGTGGCGTACCGCGCGCGTCCGCGCAAGATCCACATGA 57511

Qy 941 CGCGCCATGACGCGCATGACAGCATTTCTTCTCTCTCTCTTCTGATGTCATCTTGA 1000  
 Db 57510 CGCGCCATGACGCGCATGACAGCATTTCTTCTCTCTCTCTTCTTGA--AATCTTGA 57453

Qy 1001 TTGTCTTTTGTGTAGTCGCGCGGCTCATGTCCT 1035

Db 57452 TTGTCTTTTGTGTAGTCGCGCGGCTCATGTCCT 57418

RESULT 6  
 AK068623  
 LOCUS  
 DEFINITION  
 Oryza sativa (japonica cultivar-group) cDNA clone:J013153N23, full insert sequence.  
 AK068623  
 AK068623.1 GI:32978640  
 FLI\_CDNA; CAP trapper.  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS  
 1  
 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team,  
 Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,  
 Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,  
 Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,  
 Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ootomo,Y., Murakami,K.,  
 Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,  
 Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,  
 Nariawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nishikura,J.,  
 Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,  
 Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,  
 Kawai,J., Carninci,P., Adachi,U., Aizawa,K., Arakawa,T., Fukuda,S.,  
 Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,  
 Kagawa,I., Kondo,S., Kono,H., Miyazaki,A., Osato,N., Ota,Y.,  
 Saito,R., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shitaki,T.,  
 Yoshino,M., and Hayashizaki,Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
 Science 301 (5631), 376-379 (2003)  
 2  
 12869764  
 2 (bases 1 to 1460)  
 Adachi,U., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,  
 Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashidume,W.,  
 Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hirooka,T.,  
 Hori,F., Hotta,I., Iida,U., Iida,Y., Ikeda,R., Imamura,K.,  
 Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,  
 Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,  
 Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,  
 Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
 Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,  
 Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,  
 Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,  
 Namiki,T., Nariawa,R., Nishikura,J., Nishi,K., Nomura,K.,  
 Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,  
 Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,  
 Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,  
 Shingawa,A., Shitaki,T., Shishiki,T., Sogabe,Y., Sugano,S.,  
 Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,  
 Tagawa,A., Takahashi,F., Takaku-Akita,S., Tanaka,T., Tomaru,A.,  
 Toyota,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,  
 Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S., and  
 Yoshimura,A.  
 Direct Submision  
 Submitted (05-DEC-2001) Shoeni Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602; Japan (E-mail:ekikuchi@nias.affrc.go.jp,  
 Tel:81-29-838-7007, Fax:81-29-838-7007)  
 This clone is one of the 28k full-length cDNA clones from japonica rice.  
 URL : http://cdna01.dna.affrc.go.jp/cDNA/  
 NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

COMMENT













TITLE  
JOURNAL  
COMMENT

Messing, J., Abouelleil, A., Allen, N., Anderson, M., Anderson, S.,  
Arauchi, H.M., Bama, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S.,  
Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P.,  
Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L.,  
Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P.,  
Majors, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J.,  
Meunier, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Phunhach, P., Pierre, N.,  
Rachuga, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,  
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Submitted (20-JUN-2004) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 16, 2004 this sequence version replaced gi:48717652.  
All repeats were identified using RepeatMasker:  
Smt, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@broad.mit.edu

Bharti, AK and Messing, J: The Plant Genome Initiative at  
Rutgers, Waksman Institute, Rutgers, The State University of New  
Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA  
(http://pgr.rutgers.edu)

Butler, E and Wing, R: Arizona Genomics Institute, Biological  
Sciences West, 448A, P.O. Box 210088, University of Arizona,  
Tucson, AZ 85721, USA (http://www.genome.arizona.edu)

Project Information  
Center project name: L29613  
Center clone name: 54\_A.1

----- Consensus Information  
This consensus is derived from a shotgun assembly that has been  
manually curated. It is the best representation of the BAC that we  
can generate without further laboratory work. The draft assembly  
has been edited, and if possible, ends identified by vector as well  
as by BAC end sequences, and contigs ordered and oriented. Bases  
that are not Ns are either above Q20 or manually edited. This  
assembly was performed with phrap. All trace files for this project  
are available at the NCBI trace repository  
(http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?). An exact list of  
reads used in this assembly are available at  
http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

1 20266: contig of 20266 bp in length  
\* 20267 20366: gap of 100 bp  
\* 20367 78634: contig of 58268 bp in length  
\* 78635 78734: gap of 100 bp  
\* 78735 140318: contig of 61584 bp in length  
\* 140319 140418: gap of 100 bp  
\* 140419 198102: contig of 57684 bp in length.  
Location/Qualifiers

## FEATURES

source 1. 198102  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMBH0054A01"  
/clone\_1lb="CHORI-201 Maize B73 BAC EcoRI Library"

ORIGIN  
Query Match 10.6%; Score 140.8; DB 2; Length 198102;  
Best Local Similarity 58.5%; Pred. No. 1.1e-06;  
Matches 264; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

QY 464 CACAGCTGTTCTTCTGCTGGCTGGACAACTTGGCAAGGGGAGACTGGCGAGCATTCG 523  
DB 28743 CCAGAGCTTTCTTCTTGAAGGACATGAAGAACTAGCGGAGAGGAGCATGAGAACTTCG 28684

QY 524 CGCACTTGTGATCTCTGCGGAGCGCAAGCGAGGTGGCGAGCCAGCGGAGAACTTTC 583  
DB 28683 CGCAAGTAGTGAAGACCGCGGAGCGGAGCGAGGTGGCGAGCCAGCGGAGAACTTTC 28624

QY 584 ATCCGCTCAACTTCATGAACCGGAGCCGCGCTCCAGCATTCACGACATCACCGAGC 643  
DB 28623 ATCCGCTCAACTTCGCGGCGGAGGAGCAAGCGCGCTCCAGCATTCACGACATCACCGAGC 28564

QY 644 GTCA--CGCGCGGATCAAGTGGCCGCGAGAGGAGGCGCCCGCATCACCGGACAG 700  
DB 28563 GTCAACCTCCCGACGAGAGACCGCGGAGAGCGCGCGCTCCGCGGTACACACACACAC 28504

QY 701 GCACGCGGCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 760  
DB 28503 AACCTCTCCGTGGCTCTGTGAGACATGAACCTTCTTCCGCGCTGCGGCGGCGG 28444

QY 761 CACCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 820  
DB 28443 CACCCCTACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 28384

QY 821 CCGGTCCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 880  
DB 28383 AACGACCTCCCTCTCCGATGACAGTGTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 28324

QY 881 CAGCGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 911  
DB 28323 ACCGCTCCGTCGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 28293

RESULT 12  
AC130602/c 126532 bp DNA linear PLN 02-OCT-2004  
LOCUS DEFINITION Oryza sativa (japonica cultivar-group) chromosome 5 clone B1122D01,  
complete sequence.  
AC130602  
AC130602.5 GI:51556280  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eurharoidae; Oryzae; Oryza.  
1 (bases 1 to 126532)  
Chow, T.-Y., Hsing, Y.-I. C., Chen, C.-S., Chen, H.-H., Liu, S.-M.,  
Chao, Y.-T., Chang, S.-D., Chen, H.-C., Chen, S.-K., Chen, T.-R.,  
Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H.,  
Hsiung, J.-N., Hsu, C.-H., Huang, J.-D., Kau, P.-I., Lee, M.-C.,  
Lau, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y.,  
Yu, S.-W., Wu, H.-P. and Shaw, J.-F.  
Oryza sativa BAC B1122D01 genomic sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 126532)  
AUTHORS Chow, T.-Y. and Hsing, Y.-I. C.  
TITLE Direct Submission  
JOURNAL Submitted (13-ANG-2002) Institute of Botany, Academia Sinica, 128,  
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan

[illegible]



gene  
45797. .46030  
/gene="B1122D01.7"  
/note="unknown protein"  
45797. .46030  
/gene="B1122D01.7"  
/note="similar to rice EST AK063607"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAU90113.1"  
/db\_xref="GI:53749252"  
/translation="MLLEFVDFGFLVLLICLDYFVFLVHLIQIRABGRSTINRIGQA  
MIGPVSLSVYKLTQVYINMCIEMHGLME"  
complement(56089. .57141)  
/gene="B1122D01.8"  
/note="unknown protein"  
complement(join(56089. .56560,56666. .57141))  
/gene="B1122D01.8"  
/note="similar to rice EST AK068138"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAU90113.1"  
/db\_xref="GI:53749253"  
/translation="FMKESYMEVLPPAHYFVGQAAGGMLPDRRGGGGAWSOE  
NKYFEQALALDRDPERMERVALLIPKTYADMTYDLENDVCTIEAGLPPRY  
GAAGGGGGSTLMDGDDPDAGLFRKSCVWGGKARAGDORRKGVPWIEBHL  
FLMGLKRYGRDWMNISRNFTSRTPYQVASHAKYPIRLNSGDKDRRSIHDTTV  
NLPPDDHGNPSPSPSVLTAHSSSSAAVSEFGVAVDGPFPPLRGAGHMFMP  
HPVAQVIEAGNSHVAGGRLDSVLVQMGCGLMQPLG"

Query Match 10.4%; Score 138.6; DB 8; Length 126532;  
Best Local Similarity 59.1%; Pred. No. 2.1e-06;  
Matches 237; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Qy 466 CAGGCTGTTCTTGCTGGGCTGGACAAGTTGGCAAGGGGAGCTGGCGAGCATCTCGCG 525  
Db 56563 CAGGTTGTTCTTGATGGGGCTGAAGAAGTACGCGAGAGTACTCGAGGAACATCTCCG 56504

Qy 526 CAACCTGCTATCTCGCGGAGCGCAAGCAGAGTGGCGAGCGCGAGGAAGTACTTCAT 585  
Db 56503 CAACCTGCTGACGAGCGCGAGCGCAGAGTGGCGAGCGCGAGGAAGTACTTCAT 56444

Qy 586 CCGGCTCAATCTCAATGAACCGCGACCGCGCGCTCCAGCATCCAGCATCACCAGCT 645  
Db 56443 CCGGCTCAATCTCGCGGAGCAAGCAAGCGCGCTCCAGCATCCAGCATCACCAGCT 56384

Qy 646 CACCGCGCGAGTGGCGCGGCGAGCGAGGCGCGCGCATCACCGCGCACGAGCCAG 705  
Db 56383 CACCTCCCGACGACGACGAGCAACCCCTCGCGCTCGCGCGCGCGCTCGCTCAC 56324

Qy 706 GGGCAACCCCGCGCGCGCGCTGGGCGCGCGCGCATGAAGACCAACCAACCA 765  
Db 56323 CGGCACTCTCTCTCTCGCGCGCGCGCTCGGAGAGATTGGCGCTCGTGAAG 56264

Qy 766 CCGCGCGCGCGCGCGCGCGCATGATGAAGCGCGCGCGCATGGGCGACCCCGT 825  
Db 56263 CAAGCTTCGCGCGCGCGCGCTGGCGCGCGCGCGACCACTTCATGCCCATTC 56204

Qy 826 CGCGGCGACATGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 866  
Db 56203 CTAGCGCCAGGTGAAGATCGAGCGCGCGCATTCGCGAGCTCG 56163

RESULT 13  
AC130612/c 138467 bp DNA linear HTG 13-AUG-2002  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) chromosome 5 clone P0485G09,  
\*\*\* SEQUENCING IN PROGRESS \*\*\*, 4 ordered pieces.  
AC130612  
AC130612  
VERSION  
AC130612.1 GI:22212963  
KEYWORDS  
HTG: HTGS PHASE2.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)

REFERENCE  
AUTHORS  
1 (bases 1 to 138467)  
Chow,T.-Y., Hsing,Y.-I.C., Chen,C.-S., Chen,H.-H., Liu,S.-M.,  
Chao,Y.-T., Chang,S.-J., Chen,H.-C., Chen,S.-K., Chen,T.-R.,  
Chen,Y.-L., Cheng,C.-H., Chung,C.-I., Han,S.-Y., Heiao,S.-H.,  
Hsiung,Y.-N., Hsu,C.-H., Huang,J.-J., Kau,P.-I., Lee,M.-C.,  
Liu,H.-L., Li,Y.-F., Lin,S.-J., Lin,Y.-C., Wu,S.-W., Yu,C.-Y.,  
Yu,S.-W., Wu,H.-P., and Shaw,J.-F.  
Oryza sativa PAC P0485G09 genomic sequence  
Unpublished  
2 (bases 1 to 138467)  
Chow,T.-Y. and Hsing,Y.-I.C.  
Direct Submission  
Submitted (13-AUG-2002) Institute of Botany, Academia Sinica, 128,  
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 90592: contig of 90592 bp in length  
\* 90593 90692: gap of unknown length  
\* 90693 100409: contig of 9717 bp in length  
\* 100410 100509: gap of unknown length  
\* 100510 137114: contig of 36605 bp in length  
\* 137115 137214: gap of unknown length  
\* 137215 138467: contig of 1253 bp in length.  
Location/Qualifiers  
1. 138467  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="5"  
/clone="P0485G09"

ORIGIN

Query Match 10.4%; Score 138.6; DB 2; Length 138467;  
Best Local Similarity 59.1%; Pred. No. 2e-06;  
Matches 237; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Qy 466 CAGGCTGTTCTTGCTGGGCTGGACAAGTTGGCAAGGGGAGCTGGCGAGCATCTCGCG 525  
Db 26532 CAGGTTGTTCTTGATGGGGCTGAAGAAGTACGCGAGAGTACTCGAGGAACATCTCCG 26473

Qy 526 CAACCTGCTATCTCGCGGAGCGCAAGCAGAGTGGCGAGCGCGAGGAAGTACTTCAT 585  
Db 26472 CAACCTGCTGACGAGCGCGAGCGCGAGTGGCGAGCGCGAGGAAGTACTTCAT 26413

Qy 586 CCGGCTCAATCTCAATGAACCGCGACCGCGCGCTCCAGCATCCAGCATCACCAGCT 645  
Db 26412 CCGGCTCAATCTCGCGGAGCAAGCAAGCGCGCTCCAGCATCCAGCATCACCAGCT 26353

Qy 646 CACCGCGCGAGTGGCGCGGCGAGCGAGGCGCGCGCATCACCGCGCACGAGCCAG 705  
Db 26352 CAACCTCCCGACGACGACGAGCAACCCCTCGCGCTCGCGCGCGCGCTCGCTCAC 26293

Qy 706 GGGCAACCCCGCGCGCGCGCTGGGCGCGCGCGCATGAAGACCAACCAACCA 765  
Db 26292 CGGCACTCTCTCTCTCGCGCGCGCGCTCGGAGAGATTGGCGCTCGTGAAG 26233

Qy 766 CCGCGCGCGCGCGCGCGCGCATGATGAAGCGCGCGCGCATGGGCGACCCCGT 825  
Db 26232 CAAGCTTCGCGCGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 26173

Qy 826 CGCGGCGACATGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 866



Db 26172 CTACGCCAGGTGAAGATCGAGCGCGGCAATTCGACGTCG 26132

## RESULT 14

AY519528

LOCUS Arabidopsis thaliana 945 bp mRNA linear PLN 07-FEB-2004  
DEFINITION Arabidopsis thaliana MYB transcription factor (At1g49010) mRNA,  
complete cds.

ACCESSION AY519528  
VERSION AY519528  
KEYWORDS

SOURCE Arabidopsis thaliana (chale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 945)  
AUTHORS Qu, L. and Gu, H.  
TITLE The MYB transcription factor family in Arabidopsis: A Genome-Wide  
Cloning and Expression Pattern Analysis  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 945)  
AUTHORS Qu, L. and Gu, H.  
TITLE Direct Submission  
JOURNAL Submitted (07-JAN-2004) Life Sciences, National Laboratory of  
Protein Engineering and Plant Genetic Engineering, Peking  
University, Beijing 100871, China

## FEATURES

source

1..945  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
1..945  
/locus\_tag="At1g49010"  
1..945  
/locus\_cag="At1g49010"  
/codon\_start=1  
/product="MYB transcription factor"  
/protein\_id="AA509998.1"  
/db\_xref="GI:41618992"

/translation="MESVVAATWSREBEKAFENALALHCVBEITEDQMKNSVPSK  
ALEBKXHYQILLEDVAIENGQVPLRYHHRKGLIDEAALATSPANDSHSGSS  
BKKNPPTSGISSNGSGSRAEORRKGIPWTEBHRFLGLDFGKDWRSIS  
RNFVISRPTOVASHAOKYFRLNSMNRDRRSIHITTVNNOAPAVTGGGQPOVY  
KHRAOPORPOPOOHHPPTMAGLGMVGAAPVGQPIIAPPDHMSAVGTPTVLP  
MGTNNHHHHHLLGVAPVAVPVPPLPQHPAPSTHM"

## ORIGIN

Query Match 10.2%; Score 135.4; DB 8; Length 945;

Best Local Similarity 70.4%; Pred. No. 1.7e-05;

Matches 181; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

393 ACGACGGCGCAAGAGCTGCTCCAGGCGGAGAGAGGCGCAAGGCAATCCCATGA 452  
Db 350 ATGAGAGAAAGTGTGCTCGAGAGCTGAGCAAGAGAAAGGAAATTCATGGA 409  
453 CGAGAGAGACACAGCGTGTCTTCTGCGGGCTGGAACAAGTGGCAAGGGGGAGCTGGC 512  
Db 410 CTGAAGAGACATCGGTTGTTCTTTGGTTGGAACAAGTTGGAAAGAGATTGGA 469  
513 GGAGCATCTCGCGCAACTTCTGATCTCGCGACGCCAAGCAGGTGCGGACCGCGC 572  
Db 470 GAACATATTCAAGGAACCTTGATCTCAAGAACTCCAAACAAGTTGCAAGTATGTC 529  
573 AGAAGTACTCATCGGCTCACTTCATGAAACCGGACCGCGCGCTCCAGCATCCAG 632  
Db 530 AAAAATACTCATCAGGCTTAATCGATGAACCGAGATGAAGGGGCTGTAGCAATTCACG 589  
633 ACATACCAAGGCTCACC 649  
Db 590 ACATACCACTGTGAAC 606

## RESULT 15

AY086906

LOCUS Arabidopsis thaliana 1161 bp mRNA linear PLN 14-APR-2003  
DEFINITION Arabidopsis thaliana clone 29302 mRNA, complete sequence.

ACCESSION AY086906  
VERSION AY086906.1 GI:21405616  
KEYWORDS

SOURCE Arabidopsis thaliana (chale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1161)  
AUTHORS Haas, B.J., Volfovsky, N., Town, C.D., Troupkan, M., Alexandrov, N.,  
Feldmann, K.A., Flavell, R.B., White, O., and Salzberg, S.L.  
TITLE Full-length messenger RNA sequences greatly improve genome  
annotation  
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)  
MEDLINE 22088475  
PUBMED 12093376

REFERENCE 2 (bases 1 to 1161)  
AUTHORS Brover, V., Troupkan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
Feldmann, K.

REFERENCE 3 (bases 1 to 1161)  
AUTHORS Brover, V., Troupkan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
Feldmann, K.

COMMENT Direct Submission  
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,  
Malibu, CA 90265, USA  
This clone sequence is one of 5,000 Ceres full-length cDNAs made  
available to Tigr and Genbank. The following quality assessment of  
this set was done by comparison with known proteins: two percent  
of the clones are estimated to be 5'-truncated; less than one percent  
are 3'-truncated; approximately two percent represent alternative  
splice variants, including unspliced introns and spliced exons; one  
percent may contain premature stop codons; five percent may have  
frame shifts in a coding region. A sequence is considered to be  
5'-truncated if it lacks the translation initiation start (ATG). A  
sequence is considered to be 3'-truncated if it lacks the  
C-terminal end of the encoded protein. Please note that these cDNA  
sequences are derived from the Ws or Luer ecotypes and therefore  
may contain polymorphisms when compared to sequences from Col-0.  
GenSet carried out the library production and sequencing of the  
full-length clones. Ceres, Inc. carried out the clustering of the  
5' sequences, selection of clones, and sequence assembly.

## FEATURES

source

1..1161  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="29302"  
70..1014  
/codon\_start=1  
/product="unknown"  
/protein\_id="AA63950.1"  
/db\_xref="GI:21555865"

/translation="MESVVAATWSREBEKAFENALALHCVBEITEDQMKNSVPSK  
ALEBKXHYQILLEDVAIENGQVPLRYHHRKGLIDEAALATSPANDSHSGSS  
BKKNPPTSGISSNGSGSRAEORRKGIPWTEBHRFLGLDFGKDWRSIS  
RNFVISRPTOVASHAOKYFRLNSMNRDRRSIHITTVNNOAPAVTGGGQPOVY  
KHRAOPORPOPOOHHPPTMAGLGMVGAAPVGQPIIAPPDHMSAVGTPTVLP  
MGTNNHHHHHLLGVAPVAVPVPPLPQHPAPSTHM"

## ORIGIN

Query Match 10.2%; Score 135.4; DB 8; Length 1161;

Best Local Similarity 70.4%; Pred. No. 1.6e-05;

Matches 181; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

393 ACGACGGCGCAAGAGCTGCTCCAGGCGGAGAGAGGCGCAAGGCAATCCCATGA 452  
Db 419 ATGAGAGAAAGTGTGCTCGAGAGCTGAGCAAGAGAAAGGAAATTCATGGA 478



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 29, 2005, 21:09:58 ; Search time 4916 Seconds

(without alignments)  
3016.130 Million cell updates/sec

Title: US-10-630-636-7  
Perfect score: 1641  
Sequence: 1 MTSQAATTTTAAAAAAMTR.....GHAPYVVGVPAPAPAKMHQ 306

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh  
-Q=/cgn2\_1/USPRO.epool/US10630636/runat\_27072005\_123227\_12166/app.query.fasta\_1.455  
-DB=genemb1 -QWRT=faastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45  
-DOCALLION=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MTLEN=0 -MAXLEN=2000000000  
-USER=US10630636 @cgn\_1\_1\_4200 @runat\_27072005\_123227\_12166 -NCPu=6 -ICPU=3  
-NO MMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Genemb1:\*  
1: gb\_da:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_rts:\*  
12: gb\_by:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1641	100.0	1330	8 AY151042	AY151042 Oryza sat
2	1594	97.1	142680	8 AP004611	AP004611 Oryza sat
3	1494	91.0	1329	8 AK068565	AK068565 Oryza sat
4	857	52.2	53000	8 AP006233	AP006233 Oryza sat

5	704.5	42.9	1460	8 AK068623	AK068623 Oryza sat
6	704.5	42.9	1465	8 AK106166	AK106166 Oryza sat
7	678.5	41.3	945	8 AY151928	AY151928 Arabidops
8	678.5	41.3	1161	8 AY086596	AY086596 Arabidops
9	668.5	40.4	187154	8 OSJN00107	OSJN00107 Oryza sat
10	648.5	39.5	119091	8 AC016041	AC016041 Genomic B
11	633.5	38.6	897	8 AY151932	AY151932 Arabidops
12	633.5	38.6	928	8 AY096571	AY096571 Arabidops
13	633.5	38.6	1593	8 AY072090	AY072090 Arabidops
14	497	30.3	1988	8 AK068138	AK068138 Oryza sat
15	492	30.0	1191	8 AB083027	AB083027 Glycine m
16	489	29.8	1243	8 AB083028	AB083028 Glycine m
17	480	29.3	1795	8 AK067964	AK067964 Oryza sat
18	472.5	28.8	126532	8 AC130602	AC130602 Oryza sat
19	472.5	28.8	138467	2 AC130612	AC130612 Oryza sat
20	457.5	27.9	867	8 AY151953	AY151953 Arabidops
21	457.5	27.9	898	8 BT005473	BT005473 Arabidops
22	457.5	27.9	1640	8 AK118891	AK118891 Arabidops
23	456	27.8	892	8 AY088362	AY088362 Arabidops
24	456	27.8	915	8 AY050976	AY050976 Arabidops
25	451	27.5	679	8 AY091177	AY091177 Arabidops
26	449	27.4	897	8 AY151952	AY151952 Arabidops
27	449	27.4	928	8 BT006122	BT006122 Arabidops
28	449	27.4	1645	8 AK118135	AK118135 Arabidops
29	448	27.3	1502	8 AK111988	AK111988 Oryza sat
30	438.5	26.7	1204	8 AF239956	AF239956 Hevea bra
31	437	26.6	808	8 LES243339	LES243339 Arabidops
32	428.5	26.1	792	8 AY550308	AY550308 Arabidops
33	428.5	26.1	823	8 AY091265	AY091265 Arabidops
34	428.5	26.1	1071	8 AY056180	AY056180 Arabidops
35	428.5	26.1	1169	8 AY084487	AY084487 Arabidops
36	428	26.1	1029	8 BT012856	BT012856 Lycopersi
37	423	25.8	804	8 AY151953	AY151953 Arabidops
38	423	25.8	63898	2 OSJG00045	OSJG00045 Oryza sat
39	422	25.7	150206	8 OSJN00083	OSJN00083 Oryza sat
40	420.5	25.6	834	8 AY151931	AY151931 Arabidops
41	413	25.2	198102	2 AC145389	AC145389 Zea mays
42	410.5	25.0	79073	8 AB008271	AB008271 Arabidops
43	410.5	25.0	83485	8 ATT1153	ATT1153 Arabidops
44	403.5	24.6	1745	8 AY077454	AY077454 Antirrhin
45	399	24.3	134982	8 AP003279	AP003279 Oryza sat

#### ALIGNMENTS

RESULT 1	AY151042	1330 bp	mrna	linear	PLN 10-NOV-2002
LOCUS	AY151042				
DEFINITION	Oryza sativa (japonica cultivar-group)		transcription factor MYB51		
ACCESSION	AY151042				
VERSION	AY151042.1	GI:24850302			
KEYWORDS	mrna, complete cds.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				
REFERENCE	1 (bases 1 to 1330)				
AUTHORS	Lu, C.A., Ho, T.H., Ho, S.L. and Yu, S.M.				
TITLE	Three Novel MYB Proteins with One DNA Binding Repeat Mediate Sugar				
JOURNAL	Plant Cell 14 (8), 1963-1980 (2002)				
MEDLINE	22163024				
PUBMED	12172034				
REFERENCE	2 (bases 1 to 1330)				
AUTHORS	Lu, C.A., Ho, T.H., Ho, S.L. and Yu, S.M.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-SEP-2002) Institute of Molecular Biology, Academia				
FEATURES	Sinica, Taipei, Nankang 115, Taiwan				
SOURCE	1. 1330				
	/organism="Oryza sativa (japonica cultivar-group)"				



'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from M13 to -21M13 of the BAC clone. This sequence of OJ1005 B10 clone has an overlap with P0516004 clone (DBJ: AP003276) at the position 1 to 103,355 of 5' end. The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://xrp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES  
Source

1..142680  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="11"  
/clone="OJ1005\_B10"  
/join(3399..3512,4006..4227,4265..4328,4613..4792,6611..6996)  
/gene="OJ1005\_B10.1"  
/join(3399..3512,4006..4227,4265..4328,4613..4792,6611..6996)  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05637.1"  
/db\_xref="GI:21902090"  
/translation="MGHFLRTGTRPPSMANVAGREYQPIGRHGLARSAVRRPGRG  
GGEHREERRPRLSLGGRERKPEEVMVTPDDCSTDCNLTSPQLPLHLRSRRA  
GQGPWPRVRAHEHTCGLKPRVAGAGYFSMAGYSGHRIGVGRSLPNCR  
EPRLATLPCCHIAETADFGRTALDRNKFTYCPSESGAAGARSTARRRRTSSM  
VDGTEGVALFVVRRRRIWPRGPVADLVRPGGSGGLPEARRWIRRPRAPELL  
SRPPRAAHSRVPPEPELAAPFAHSVSGRGRCGCFIQP"  
/join(9727..9823,11568..11674)  
/gene="OJ1005\_B10.2"  
/join(9727..9823,11568..11674)  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05638.1"  
/db\_xref="GI:21902091"  
/translation="MAEAACTRASQOQDEAAGIENGCKQLPEDTTCVAAAGAPP  
SMSNSVALINMGCGSSLPHEB"  
/complement(13538..14391)  
/gene="OJ1005\_B10.3"  
/complement(13538..14391)  
/note="gag-pol polyprotein  
probably inactive due to frameshift(s) in CDS  
probably inactive due to stop codon(s) in CDS  
pseudogene"  
/pseudo  
/join(16900..17014,17370..17504,18435..18718)  
/gene="OJ1005\_B10.4"  
/join(16900..17014,17370..17504,18435..18718)  
/gene="OJ1005\_B10.4"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05639.1"  
/db\_xref="GI:21902092"  
/translation="WVTSKYGHDSGBELVILGPHPTATELLOLPILPTGSGIR  
KGRKAHNSIMLTASYVTKKALQWQHLPIISCHKGCDOLGSGSGCGRRRAGC  
PEKAVTTAGDSEBGERLALQVALHAKGAEGEKRETRREGEDLTKKRRVQERT  
ATWTLRAKRTSGSDN"  
/join(22941..24398,24894..25130)  
/gene="OJ1005\_B10.5"  
/join(22941..24398,24894..25130)  
/gene="OJ1005\_B10.5"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05640.1"  
/db\_xref="GI:21902093"

/translation="MTIFQTLTWMAAALVACIFISFRLLKODYVDPKDKODHK  
NIRSLNVFYGLALAHSMYSYIILVAQPLVAPVRVHVSYSKLHDMQMDVSEYAK  
KLAFIRNVREALEMMULIFGKELASDSIDPOLKVRILDIYIIRMEKHLSPAGYL  
DIYIEYKRGKRVLSIRASTETLERVNMGLNKSLEEBTRHAASIVELA.PYLL  
FENLPAMPQILASLLTGRERINRPNPDINSVELTCGVKILBELVNPDRRSVADA  
NVLAKTVELNFRDCKVPVDETERNSOEIVASLNLVHLKVTGTGEELARST  
ISKNSIEMSNIRKILYQHDNKSLSVAHVILSCLMHETARAVIGSSCQIVKLVS  
SVLRPMPIYODGNNGSTVADSAQAVALSTGSEYNRATILDIINLELVGMLFAS  
MEGRIMLHLKLRTYSGPGYQNLKIDGSLPKDLKELEKSGKLDGFIGLSSQ  
ICHLQALDPKALQSANLMDTYAODGTEVVOISYHRFPDGKSKNTGSDMAQK"  
/complement(join(26516..26605,26825..26971))  
/gene="OJ1005\_B10.6"  
/complement(join(26516..26605,26825..26971))  
/gene="OJ1005\_B10.6"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05641.1"  
/db\_xref="GI:21902094"  
/translation="MENIYTSRVLPKYLGGDRVVISYKIQCLRVLLGDQDDPWYKR  
DPQGRILVPYVNTLIIRYSITRQLYTPQTCY"  
/complement(join(28040..28278,28339..28375))  
/gene="OJ1005\_B10.7"  
/complement(join(28040..28278,28339..28375))  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05642.1"  
/db\_xref="GI:21902095"  
/translation="WGAADPAETAAHPABSGVFSRRPRGQREGRVAAAVRRHG  
GSGSVGRSGRRIRERRATTAAVAVFORLDSPTARTATDLC"  
/join(30138..30218,31910..32079,32238..32400)  
/gene="OJ1005\_B10.8"  
/join(30138..30218,31910..32079,32238..32400)  
/gene="OJ1005\_B10.8"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05643.1"  
/db\_xref="GI:21902096"  
/translation="MQLNPKERTRNKTKAKGRCARIVLNKNAATTIHDVGVGCG  
GGGTRDLPRLMRSTRTATYAPERSHADCKYDGGCGGCRADNGDGLTAVAPAFTE  
VVDGGSYGGGRHMRERKREKRGGEDEDTDS"  
/33576..33806  
/gene="OJ1005\_B10.9"  
/33576..33806  
/gene="OJ1005\_B10.9"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05644.1"  
/db\_xref="GI:21902097"  
/translation="MRRVTQMKSTAAAGDAATEGDEDDDEDDGVEDDEDD  
DGGCGHGBAERGDDAAAGAEVARRGGG"  
/join(35782..36026,36318..36486)  
/gene="OJ1005\_B10.10"  
/join(35782..36026,36318..36486)  
/gene="OJ1005\_B10.10"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05645.1"  
/db\_xref="GI:21902098"  
/translation="MAVPPKNDAAFPDPPNDERLPEKXLTATVAVRIAPFYVWGC  
GALAFNATVILGGYATLILKODFMFTIIVFEATCPGKPTLTSRLIADTPLDSD  
NICVRGYSRRSPFTAGTDLKLFGMRHSHREL"  
/complement(join(39042..39337,40177..40346,40543..40590,41257..41417,42100..42248))  
/gene="OJ1005\_B10.11"  
/complement(join(39042..39337,40177..40346,40543..40590,41257..41417,42100..42248))  
/gene="OJ1005\_B10.11"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAC05646.1"  
/db\_xref="GI:21902099"  
/translation="MARTGAAVAHARTDGLATFPMTGKKYELGCTSSDEPPRAH

[illegible]

Db	139958	AAGGCGCGCCGACCATGGGCGCACCCCGTCGCGGCGCACATGGTGCCCGCGCGTGGGAC	139917
Oy	280	rProValValPheProProGlyHisAlaProTyValValProValGlyTyrProAlaIar	300
Db	139918	GCCCGTGGTGTTCCTCCGCGGGCGCACGCCCGTACCTGTGCGCCGTCGGGTACCCGCGCC	139977
Oy	300	oProAlaLysMetHisGln	306
Db	139978	ACCGGCCCAAGTGCACCA	139996
RESULT 3			
AK068565			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group)	cdna clone:J013154G12, full	
ACCESSION	AK068565		
VERSION	AK068565.1	GI:32978582	
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE			
AUTHORS	1 The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,S., Yahagi,M., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Oono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Naitakawa,R., Sugiyama,A., Mizuno,K., Kobayashi,M., Xie,Q., Lu,M., Ikeda,T., Ishibiki,J., Kawamata,M., Yoshimura,H., Mura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Haenlidume,W., Hayatsu,N., Imotani,K., Ishi,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Oosato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science	301 (5631), 376-379	(2003)
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2 (bases 1 to 1329)		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hiroane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,T., Kusumegi,T., Li,C., Lu,M., Koya,S., Kurihara,C., Kurotaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Mura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murakami,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niihara,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Oono,M., Ohtsuki,K., Oka,M., Ooka,H., Oosato,N., Ota,Y., Oono,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Kubira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,M., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S., and Yoshimura,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-DEC-2001)	Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamadai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,	



ORIGIN		/db_xref="taxon:39947" /chromosome="1" /clone="B1249E06"	
Alignment Scores:			
Pred. No.:	6.45e-27	Length:	53000
Score:	857.00	Matches:	157
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.22%	Indels:	0
DB:	8	Gaps:	0
US-10-630-636-7 (1-306) x AP006233 (1-53000)			
Qy	150	Argleupheuleuaglyleuaspheglylyspthyspripargserileserarg	169
Db	63	AGCGTGTCTTGGCTGGGCTGACAGTGGCGCAAGGGGAGCTGGCGAGCATCTCGCGC	122
Qy	170	Asnphvalliesserargthprctthglinalaserhisalaglnlytyrphelle	189
Db	123	AACTTCGTCACTTCGCGGACCGCAACGACGCTGGCGACCGCGCAGAACTCTCATC	182
Qy	190	Argleuphanserleasnarargapargargserierlehisaspriethserval	209
Db	183	CGCTCAACTCCAGAACCGGACCGCGCTTCGACATCCAGACATCCACGCGTC	242
Qy	210	Thralaglyaspglinalalaglinalalproliethrthlyshlinalathr	229
Db	243	ACCGCGCGCATCAAGTCCGCGCCAGAGGGCCCGCGATCACCAGCCAGCCACG	302
Qy	230	Glyanproalaalalaleuaglyprobglymetlyshishishishishis	249
Db	303	GGCAACCCCGCGCGCGCGCTGGCGCGCGCGCATGAAGCACACACACACACAC	362
Qy	250	Proglyglyalaproproprometprmetprserialaaprometglyhisproval	269
Db	363	CGGGCGCGCGCGCGCGCGCGCATCGATGACGCGCGCGCATGGCGCACCGCGTC	422
Qy	270	Alaaglyshmetvalproalaalalalvalgltyrprovalalpheproproglyhisala	289
Db	423	GCGCGCCACATGGTGCCCGCGCGCGCTGGCGGACGCGGTGTGTCCCGCGGACCGG	482
Qy	290	Protyrvalvalprovalgltyrproalaproproalalyshmetshl	306
Db	483	CCGTACGTCGTGCGCGCTGCGCTACCGCGCGCACCGCCAGATGACCA	533
RESULT 5			
AK068623			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1			
The Rice Full-length cDNA Consortium, National Institute of			
Agrobiological Sciences Rice Full-length cDNA Project Team:			
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,			
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotte, I.,			
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,			
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International			
Science Genome Sequencing & Analysis Group, Ohtsuki, Y.,			
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,			
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,			
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishikawa, J.,			
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J.,			
Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
1			
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,			
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,			
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oota, N., Ota, Y.,			
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,			
Yoshino, M., and Hayashizaki, Y.			
Collection, mapping, and annotation of over 28,000 cDNA clones from			
Japanese rice			
Science 301 (5631), 376-379 (2003)			
2			
(bases 1 to 1460)			
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,			
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,			
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T.,			
Hori, F., Hotte, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,			
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,			
Kagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,			
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,			
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,			
Koya, S., Kurihara, C., Kurosaki, T., Kusunegi, T., Li, C., Lu, M.,			
Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A.,			
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,			
Namiki, T., Narikawa, R., Nishikawa, J., Nishi, K., Nomura, K.,			
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,			
Osato, N., Ota, Y., Ohtsuki, R., Ryu, R., Satoh, H., Sakai, C., Sakai, K.,			
Sekazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,			
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,			
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,			
Tagawa, A., Takahashi, F., Takaku-Akishi, S., Tanaka, T., Tomaru, A.,			
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,			
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and			
Yoshimura, A.			
Direct Submission			
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of			
Agrobiological Sciences, Department of Molecular Genetics, Head of			
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki			
305-8602, Japan (E-mail: sikuchi@nias.affrc.go.jp,			
Tel:81-29-838-7007, Fax:81-29-838-7007)			
This clone is one of the 28k full-length cDNA clones from japonica			
rice.			
URL: http://cdna01.dna.affrc.go.jp/cDNA/			
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,			
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,			
Ishikawa, M., Yamada, H., Ooka, H., Hotte, I., Kojima, K., Namiki, T.,			
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and			
Yamamoto, M.			
PAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,			
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,			
Kodama, T., Kurosaki, T., Kusunegi, T., Lu, M., Masuda, H., Mura, J.,			
Mizuno, K., Narikawa, R., Nishikawa, R., Nishikawa, N., Sugano, S.,			
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,			
Yoshimura, A., Matsubara, K., and Murakami, K.			
Genome Exploration Research Group in Riken Genomic Sciences Center			
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,			
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,			
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,			
Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,			
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,			
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,			
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Mura, J.,			
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,			
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,			
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,			
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,			
Takaku-Akishi, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,			
Yasunishi, A., and Hayashizaki, Y.			
Location/Qualifiers			
1. 1460			
/organism="Oryza sativa (japonica cultivar-group)"			
/mol_type="mRNA"			
/cultiivar="Nipponbare"			
/db_xref="taxon:39947"			
/clone="U01315N23"			
FEATURES			
source			





URL : <http://cdna01.dna.afrc.go.jp/cDNA/>  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Onsenki, K., Shishiki, T. and Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Otsomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Aikawa, T., Arikawa, T., Carninci, P., Fukuda, S., Hanganaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa, Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Koude, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami, T., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yeanishi, A. and Hayashizaki, Y.  
 location/Qualifiers  
 1. 1465  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultiVar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="001-208-B07"

## ORIGIN

Alignment Scores:  
 Pred. No.: 8.4e-22 Length: 1465  
 Score: 704.50 Matches: 166  
 Percent Similarity: 58.57% Conservative: 22  
 Best Local Similarity: 51.71% Mismatches: 76  
 Query Match: 42.93% Indels: 57  
 DB: 8 Gaps: 13

US-10-630-636-7 (1-306) x AK106166 (1-1465)

QY 12 ALaAlaAlaAlaAlaAlaTPrThArGluAaSpAaPlySaAlaPheGluAaAlaLeuAla 31  
 Db 279 TCCCGTTCGGCGGGTGGTGAACCGCAGCAGAGAAAGCGCTTCGAAACCCCTGGCG 338  
 QY 32 ALaCySaAlaAlaProProAlaProAlaSpGlyAlaProAaPaPaPaPaPaPaPaPa 51  
 Db 339 ACGGTG-----GGGATGACACAGAGAGAAAGCGACGGCTTGTGGAGAG 383  
 QY 52 LeuAlaAlaSerValProGlyAlaAaGSerAlaGluGluValAaGArgHisTyrGluAla 71  
 Db 384 CTACGGAGAGCGCGTGAAGGGG---AAGACGGCCCAAGCGAGTGAAGCGCACTTACGAGTGG 440  
 QY 72 LeuValGluAaPValAlaAlaAlaIleAaPAlaGlyAaGValProLeuProAaGTYrAlaGly 91  
 Db 441 CTGGTGGAGAGACGTGACACGCGATCGAGCGCGGGCGGTGGCGCTCTGTGTACGCCGGC 500  
 QY 92 GluGluSerAlaAlaProProAaPAlaA-----GlyAlaAlaAlaAlaAlaSer 108  
 Db 501 -----GACGGGGCGCTGACAGAGAGCGCTCTGGCGGAGAGTGGAGAG 539  
 QY 109 LysAaPGLyGlyHisArgAaGAspGluAaGlyGlyGlyGlyGlyTyrAaPGLyGly 128  
 Db 540 AAGGGGGGTGT-----GGGGAGAGAGGTGAGGTGAGGGGGG 578  
 QY 129 Lys-----SerCySerTylAaGluGluGluGluAaGArgAaGlyGlyIleProTyr 144  
 Db 579 CATGGGAGAAAGGGGTGGCTTAAGTCTCTGTAGAGAGAGCGCCGGAAAGGGATGCTGG 638

QY 145 ThGluGluGluHisArgLeuPheLeuLeuGlyLeuAaPlyPheGlyTyrGlyAaPTrp 164  
 Db 639 ACCGAGAGCAGACAGACGGCGTTCCTCTTGGACTTGAAGAAAGTACGCAAGCGACTGG 698  
 QY 165 ArgSerIleSerArgAaPheValIleSerArgThrProThrGluValAlaSerHisAla 184  
 Db 699 AGAGATCTCCAGAAACTTGTGATCTCAAGACACCCACCAAGTACTATCATGCA 758  
 QY 185 GlnLysTyrPheIleArgLeuAaSerMetAaAaArgAaPArgAaGSerSerIleHis 204  
 Db 759 CAGAGATTTTATTCCTCGCTGAATCTCAATGAACAGAGAGCGCGCATCAAGTATCAT 818  
 QY 205 AspIleThrSerValThrAlaGlyAaPArgValAlaAlaGluGlnGlyAlaProIleThr 224  
 Db 819 GACATACCCGCTGAAACCAATGAGATACATCTGCTCTAG-----GGCCCAATCACA 872  
 QY 225 GlnHisGlnAlaThrGluAaPProAlaAlaAlaAlaLeuGlyProProGlyMetLysHis 244  
 Db 873 GGT---CAGCCAAATGCG---CATCAGCA-----AATCTGGAATAATCTCT 914  
 QY 245 HisHisHisHisHisProGlyGlyAlaProProProweProweTyrSerAlaAlaPro 264  
 Db 915 AAGCAGTCTTACAGCCAGCAAAATGGCGCTCCAGCGCTCATGCTTATGTACGACA--- 971  
 QY 265 MetGlyHisProValAlaGlyHisMetValProAlaAlaValAlaGlyThrProValAlaPhe 284  
 Db 972 ATTGACAGCCAAATGTGTGTCTCTTGTG---TCCGCAATGGACATCTCTGTACACTT 1028  
 QY 285 ProProGlyHisAlaProTyrValVal-----ProValGlyTyrProAlaProProAlaLysMet 293  
 Db 1029 CCGTCTCTGCTGACCTCATATAGCTTATGACATGATGATGATGATGATGATGATGATGAT 1088  
 QY 294 -----ProValGlyTyrProAlaProProAlaLysMet 304  
 Db 1089 GTCCCTGTGCCCCAGTAAACATGCTCCATGCTCAATGCCCTACATGCCGACCAACATCT 1148  
 QY 305 His 305  
 Db 1149 CAT 1151

RESULT 7  
 AY519528  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana MYB transcription factor (At1g49010) mRNA,  
 complete cde.  
 ACCESSION  
 AY519528  
 VERSION  
 AY519528.1 GI:41618991  
 SOURCE  
 Arabidopsis thaliana (thale cress)  
 ORGANISM  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
 1 (bases 1 to 945)  
 AUTHORS  
 Qu, L. and Gu, H.  
 TITLE  
 The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide  
 Cloning and Expression Pattern Analysis  
 JOURNAL  
 Unpublished  
 REFERENCE  
 2 (bases 1 to 945)  
 Qu, L. and Gu, H.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (07-JUN-2004) Life Sciences, National Laboratory of  
 Protein Engineering and Plant Genetic Engineering, Peking  
 University, Beijing 100871, China  
 FEATURES  
 location/Qualifiers  
 1. 945  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 1. 945  
 /locus\_tag="At1g49010"  
 1. 945  
 /locus\_tag="At1g49010"







Oy	118	ArglysglygylgylgylTyrAaspGlyIylysSerCysSerIysAlaIugInGln	137
Db	58186	AAGACGGAGCGCGCGGCAGGTATGACAGCGCGGCACAAGACTGTCTCCAAAGCGAGCAAGAG	58127
Oy	138	ArGaArgIysgIyLleProTrpThGlUGluHis-----	149
Db	58126	AGCGCGCAAGGGCATCCCATGACCGAGGAAGAGACAGTAAGTATACACGCCGCAATTCC	58067
Oy	149	-----	149
Db	58066	GGAATACAGACGACGAGTTGTGTTCGTCTTCTTCCTCGGAACACAGCTCGCTCTCTG	58007
Oy	149	-----	149
Db	58006	CTGCATCTGCTGCGTGACGCAAGCGAAATCCATGCGGTGATGTTCATGCGATTGCGATTGG	57947
Oy	150	-----ArgLeuPheLeuLeuGlyLeuAspIysPheGlyIysgIyAspTrpArgSer	166
Db	57946	ATTGTGGCGCAGGCGTCTTCTTCTCTGGGCTGGACAGTTCCGGCAAAGGAGACTTGCACAGC	57887
Oy	167	IleSerArganPheValIleSerArgThrProThrglInvalAlaserHisAlaGlnIys	186
Db	57886	ATCTCGGCAACTTCGTCATCTCGCGAGCGCACGAGGTGGCGAGCCACGCGCAGAAG	57822
Oy	187	TyrPheIleatgLeuAnserMetAsnArgAspArgArgSerSerIleHisapIle	206
Db	57826	TACTTCATACCCCTCCAACCTCCATAMACGGGATGTCGCGCTCCACAGCATCCACGACATC	57767
Oy	207	ThisSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrgIyHis	226
Db	57766	ACCAGCGTCACCGCGCGGAGAC--GTGGCGCGCAGACAGGAGC---CCGATCACCAGC---	57716
Oy	227	GlnAlaThrgIyAsnProAlaAlaAlaAlaLeuGlyProProGlyMetIySHisHis	246
Db	57715	CAGGCGAA-GGCAGCCCT--GCGCGCGCGCTCGGCGCAAGGGCATGAAG-----	57665
Oy	247	HisHisIasProGlyGlyAlaProProPrometProMetIySerAlaAlaPrometGly	266
Db	57668	-----CACCGCGCGCGG-----CCGCGCATGGCCATGTACGATGGCGCGCCACATGGCG	57622
Oy	267	HisProValAlaGlyHisMetValProAlaAlaValGlyThrProValValPheProPro	286
Db	57620	CACCCCGTCGCGGCGCACATGTGCTCTCGCGCGCTCGGACGCGCGGATGTCTCCGCGG	57567
Oy	287	GlyHisAlaProTyValValProValGlyTyxProAlaProProAlaIyMetHisGln	306
Db	57560	GC-----TACC GGCGCGCCCTCGGCAAGATGACACAA	57522
RESULT 10			
AC016041/c	119091 bp	DNA	linear PLN 11-OCT-2000
LOCUS			
DEFINITION	Genomic sequence for Arabidopsis thaliana BAC F2J15 from chromosome I, complete sequence.		
VERSION	AC016041		
KEYWORDS	AC016041.5 GI:6910416		
SOURCE	HTG.		
ORGANISM	Arabidopsis thaliana (chale cress)		
REFERENCE	Arabidopsis thaliana		
AUTHORS	Eukariotae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 119091)		
	Shim, P., Brooke, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Alatafi, H., Bei, Q., Chin, C., Chlou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howling, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thayer, A., Toriumi, M., Vayberg, M., Yu, G., Federick, N. A., Theologis, A. and Becker, J. R.		
TITLE	Genomic sequence for Arabidopsis thaliana BAC F2J15 from chromosome I		
JOURNAL	unpublished		

REFERENCE	TITLE	JOURNAL	AUTHORS
2 (bases 1 to 119091)	Direct Submission	Submitted (20-NOV-1999)	Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
3 (bases 1 to 119091)	Direct Submission	Submitted (06-FEB-2000)	Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
4 (bases 1 to 119091)	Direct Submission	Submitted (12-MAY-2000)	Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
5 (bases 1 to 119091)	Direct Submission	Submitted (11-OCT-2000)	Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinm,P., Altafi,H., Bel,B., Chin,C., Chio,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukhareky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federpiel,N., Theologis,A. and Ecker,J.	Direct Submission	Submitted (12-MAY-2000)	Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinm,P., Altafi,H., Bel,B., Chin,C., Chio,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukhareky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federpiel,N., Theologis,A. and Ecker,J.	Direct Submission	Submitted (11-OCT-2000)	Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
On Feb 6, 2000 this sequence version replaced gi:6684171.	Location/Qualifiers		
1..119091			
/organism="Arabidopsis thaliana"			
/mol_type="genomic DNA"			
/db_xref="taxon:3702"			
/chromosome="1"			
/clone="F2J015"			
complement(join(1671..1736,1844..2457,2554..2947,3020..3079))			
/note="putative actin 11 gp P53496; similar to ESTs gb N96401.1, gb N95917.1, gb T45783.1, gb H77213.1, gb R90244.1, gb AA712648.1, gb N38151.1, gb A1352754.1, gb H76490.1, gb T21319.1, gb T46255.1, gb AA042602.1, gb H85824.1, gb H76772.1, gb H76502.1, gb R90429.1, gb T88401.1, gb N38636.1, gb T20894.1, gb T43293.1, gb AA712760.1, dbj AB038718.1"			
/codon_start=1			
/evidence=not_experimental			
/product="F2J015.1"			
/protein_id="AA069724.1"			
/db_xref="GI:7770354"			
/translation="MADADDIOPICDNGHGMVAKGAPDAPRAVPSVGRPRHGGVGMGNORDAYVEDSAOSKRGITLTPKRGHGVKASVNDMDKEXWHHTFVLEIAPBEHPVLTEAPLNPKAKREKMTQIMETFRNSPAMVALQAVLSLASGETTIVLSDGGVSHVAPLYEGFSLEPHALILDLADGRDITDVMILTEKRGWFTTARETRVRIKELUSFVAVDEEMETSSKSLSEKYNLEKQVLTIGERFPCPEVLFPSPSVEGMAAGIHEHYVNSIMKCDVDIRDLGNYIVLISGCTGTGSGIDRMSKEKITALAPSMTKIKVAVPERKSWIGGSIASLSTFOOMWISKAEXEDAGPGIVHRKCF"			
complement(join(9639..9780,9617..10114,11072..11183,12944..13672))			
/note="putative RING zinc finger protein gb AA16660.1; similar to gb N17302.1, gb T75752.1"			
/codon_start=1			
/evidence=not_experimental			
/product="F2J015.2"			
/protein_id="AA069723.1"			





Db 86508 GGGATCCATGACTGAAGAAGCATGGTTCTTCTTCTTCTTCATCTCCATCAT 86449  
Qy 149 ----- 149  
Db 86448 CTATCATTTGATATTATTAGACTTAAAGATGAAAAAATTTGGAAGATCTTTACC 86389  
Qy 150 ----- 160  
Db 86388 TTTATTGAATCTGTTGTTGTTGGAAGGTTGTTCTTTGGGTTTGAACAATTTGGG 86329  
Qy 161 LysGlyAspSerIlePargSerIleSerArgAspPheValIleSerArgThrProThrGlnVal 180  
Db 86328 AAAAGAGTTTGAAGAAGCATTTTCAAGAACTTTTGATCTTCAAGAACTCCAAACAGATT 86269  
Qy 181 AlaSerHisAlaGlnLeuTyrPheIleArgLeuAsnSerMetAsnArgAspArgArg 200  
Db 86268 GCAAGTCATGCTCAAAAATACCTTCATCAGCTTAATCGATGAACCGAGATAGAAGCGG 86209  
Qy 201 SerSerIleHisAspIleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGly 220  
Db 86208 TCTAGCATTCACGACATCAC-----ACTGTGAACAATCAAGCTCTCGGTTACAGGA 86155  
Qy 221 AlaProIleThrGlnIleGlnAlaThrGlyAsnProAlaIleAlaIleGlnGlyProPro 240  
Db 86154 GGAGGACAACAACCGCAAGTGGTTAAACATAGACCAAGCTACGCCAACACCAACCCACAA 86095  
Qy 241 GlyMetLysHisHisHisHisHisHisProGlyGlyAlaProProProMetPro----- 258  
Db 86094 CCGCAACCAACAACATCAT-----CCCCCAACAATGGCTGGATTA 86053  
Qy 259 ---MetTyrSerAlaAlaProMetGlyHisProValAlaGly-----HisMetVal 274  
Db 86052 GGGATGATGGTGTGCGCCAGTGGACACCATGATCGACCAACCTGATCATATG--- 85996  
Qy 275 ProAlaAlaValGlyThrProValValPheProPro----- 286  
Db 85995 GGTTCAGCTGTTGGAACCCCTGATGCTTCCACTCCAAATGGCAACTCATCATCAC 85936  
Qy 287 -----GlyHisAlaProTyrValValProValGlyTyrProAlaProPro 301  
Db 85935 CATCACCATCATTTGGAGTTGCTCTTATGCTGTACGG---GCTTATCCGGTACCGCA 85879

RESULT 11  
AY519532  
LOCUS AY519532 897 bp mRNA linear PLN 07-FEB-2004  
DEFINITION Arabidopsis thaliana MYB transcription factor (At5g08520) mRNA,  
complete cds.  
ACCESSION AY519532  
VERSION AY519532.1 GI:41619007  
KEYWORDS Arabidopsis thaliana (chale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 897)  
AUTHORS Qu,L. and Gu,H.  
TITLE The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide  
Cloning and Expression Pattern Analysis  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 897)  
AUTHORS Qu,L. and Gu,H.  
TITLE Direct Submission  
JOURNAL Submitted (07-JAN-2004) Life Sciences, National Laboratory of  
Protein Engineering and Plant Genetic Engineering, Peking  
University, Beijing 100871, China  
FEATURES  
SOURCE  
1. .897  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
1. .897  
/locus\_tag="At5g08520"  
gene

CDS  
1. .897  
/locus\_tag="At5g08520"  
/codon\_start=1  
/product="MYB transcription factor"  
/protein\_id="AA510002.1"  
/db\_xref="GI:41619008"  
/translation="MTVEKSDGSVWSREDDIAFERALANTDESEREKIAADPG  
KSYEQIKENYELVEDVTRLESQVPLPAFGSPGSGNAGHAGDEGASKGSGNHADES  
NQAGKSKSDERRKGIAMTEDEHRLFLGLDKYKGDWRSISNFPVTRPTGVASHA  
OKYFIRLNSMKNDRRSIHDIIVGNADVSTPOGPIITGQNNNNNNNNNNNNSSPAY  
AGGNLSAKQAVSCAPPGPMVGTPIAGQAVGTPTVNLPAIPHMAVGVAAPVGSIV  
PGANNINIGMPYTPMPTPTAHR"  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,53e-19 Length: 897  
Score: 633.50 Matches: 155  
Percent Similarity: 58.60% Conservative: 29  
Best Local Similarity: 49.36% Mismatches: 81  
Query Match: 38.60% Indels: 49  
DB: 8 Gaps: 14  
US-10-630-636-7 (1-306) x AY519532 (1-897)  
Qy 8 ThrThrThrThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 27  
Db 4 ACCGTGAGGAGAGTTAGTATGATGTTCTGTGTGAGATGAGGAGATGATATTTGCTTTGAG 63  
Qy 28 AsnAlaLeuAlaAlaCysAlaAlaProProProAlaAspGlyGlyAlaProAspAspAsp 47  
Db 64 AGAGCTTA-----GCCAATATATCCGATGAATCAGAGAA 99  
Qy 48 TRPHeAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGlnGlnValAlaArgArg 67  
Db 100 CGGTGGGAGAAAGATTGCTGCAGACGTTCCAGGC---AAAGTGTGGAACAGATTAAGAA 156  
Qy 68 HisTyrGlnAlaLeuValGluAspValAlaAlaIleAspAlaGlyArgValProLeuPro 87  
Db 157 CATTACAGAGCTTTTAAAGTTGAAGATGTTACTAGGATGAATCAGAGATGTGTGCTCTTCTCT 216  
Qy 88 ArgTyrAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 106  
Db 217 GCCTATGGTCTCTCTGAAGATCG-----AATGGCCATCTGCTGTAAGAGCA 267  
Qy 107 AlaSerLysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGlyGlyTyrAsp 126  
Db 268 AGTAGTAAGAAAGAGGTAAACAGTCAT-----CGGGAGAGCTTAACCA 312  
Qy 127 GlyGlyLysSerCysSerLysAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 146  
Db 313 GCAAGTAA-----TCMAAGTCCGATCAAGAACGAAAGGGTATCGCGTGAACGAA 366  
Qy 147 GlnGlnHisArgLeuPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 166  
Db 367 GATGAGCAGCGTTATTTCTTTGTTGGTGAATGATGACGGAAAGGTATTTGGCTAGC 426  
Qy 167 IleSerArgAspPheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLeu 186  
Db 427 ATTTCTGCAACTTTGTAATGAACAGAACACCGAACCAAGTTGCGAGCATGCTCAAAAG 486  
Qy 187 TyrPheIleArgLeuAsnSerMetAsnArgAspArgArgArgSerIleHisAspIle 206  
Db 487 TATTTCAATGCTGTTAAATTAATGAACAAAGACAGAGGAGATCAATTCACGACATC 546  
Qy 207 ThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGlnHis 226  
Db 547 ACTAGTGTGGCAACGCAGAT---GTCTCAACCCCAAGGA---CCAAATCACTGGTCAAG 600  
Qy 227 GlnAlaThrGlyAsn-----ProAlaAlaAlaAla 236  
Db 601 AACACAGCAATTAACAACAACAACAACAACAACAACAAGTTCTCTGCTGTGCTGGA 660  
Qy 237 LeuGlyProProGlyMetLysHisHisHisHisHisHisProGlyGlyAlaProProPro 256



DB	661	GGAGGAACAAATATGACCGAAG-----CAAGCGCTCTCCAGCAACACTCTGA	708
Oy	257	MetProMetYrYrSerAlaAlaProMetGlyHisProValAlaGlyHisMetValProAla	276
Db	709	CCTCTTAATGATATGGAACAACCGCCATAGATGTCAGCA-----	744
Oy	277	AlaValAlaGlyThrProValValPhePro---ProGlyHisAlaProTyrValVal-----	293
Db	745	GCAGTGTGAACACCGATACCTCCCACTCCACTCCACTCAGTCATGAGTTATGAGATTATGTCG	804
Oy	294	---ProVal---GlyTyrProAlaProProAlaValMetHis	305
Db	805	GCTCCAGTCCCTCGCTCAGTGTCTTCTGTGTCAGCAATGAC	846
RESULT 12			
LOCUS	AY096571	928 bp	mRNA linear
DEFINITION	Arabidopsis thaliana unknown protein (At5g08520)	mrna, complete	
ACCESSION	AY096571		
VERSION	AY096571.1		
KEYWORDS	GI:20465998		
SOURCE	FLI CDNA.		
ORGANISM	Arabidopsis thaliana (chale cress)		
REFERENCE	Arabidopsis thaliana		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
	1 (bases 1 to 928)		
	Yamada, K., Banh, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M., Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C., Tortum, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R. W., Ecker, J. R. and Theologis, A.		
JOURNAL	Arabidopsis Open Reading Frame (ORF) Clones		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 928)		
	Yamada, K., Banh, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M., Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C., Tortum, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R. W., Ecker, J. R. and Theologis, A.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (16-Apr-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, NY 12242, USA		
AUTHORS	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA': Seki, M., Narusaka, M., Ishida, J., Saito, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
TITLE	The Salix, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Banh, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M., Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C., Tortum, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R. W., Ecker, J. R. and Theologis, A.		
COMMENT			
	Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.		
	Location/Qualifiers		

[illegible]

Qy	187	TyrPheIleAglLeuAsnSerMetCysAspArgAspGlySerSerIleHisAspIle	206
Db	487	TATTTCATTTCGTCTTAATTCANTGACAAAGACAGAGCCGATCAAGATTCACACATC	546
Qy	207	ThiSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGlyHis	226
Db	547	ACTAGTGTGGCAACGACGAGAT---GTCTCAACGCCACAAGGA---CAATCAGTGGTCAG	600
Qy	227	GlnAlaThrGlyAsn-----ProAlaAlaAlaAla	236
Db	601	AACAACAGCAATTAACAAACAACAACAACAACAACAGTTCTCTCGTGGTCTGGA	660
Qy	237	LeuGlyProProGlyMetLeuHisHisHisHisHisHisHisProGlyGlyAlaProProPro	256
Db	661	GGAGGAAACAATCAGCCACG-----CAAGCCGTCTTCAAGCAACCACTTGGGA	708
Qy	257	MetProMetTyrSerAlaAlaProMetGlyHisProValAlaGlyHisMetValProAla	276
Db	709	CCTCTTAATGTATGGAAACACCCGCCATAGCTACGCCA-----	744
Qy	277	AlaValGlyThrProValValPhePro---ProGlyHisAlaProTyrValVal-----	293
Db	745	GCAAGTGGAAACACGATCACTCCCAAGTCCCACTCAACATGAGGCTTATGAGTTATCGG	804
Qy	294	---ProVal---GlyTyrProAlaProProAlaLysMetHis	305
Db	805	GCTCCAGTCCCTGGCTCAGTGGTCTCTGGTGAAGCAATGAAC	846

Db	760	GCTATGGGCTCTCTGAAGATCG-----AATGCCATCTGCTGATGTAAGAGACA	810
Qy	107	AlaSerLySAspGlyGlyHisArgArgAspGluArgLyGlyGlyGlyTyrAsp	126
Db	811	AGTACTAAGAAAGAGGTACAGTCAT-----GGGGAGAGCTTACCA	855
Qy	127	GlyLyLySerCysSerLySAlaGluGlnIuaArgLyGlyIleProTyrThrGlu	146
Db	856	GCAGGTAA-----TCAAGTCCGATCAGAAACAGCAAGGATGCGCGTGCAGACA	909
Qy	147	GluGluHisArgLeuPheLeuGluGlyLeuAspLyPheGlyLyAspTyrArgSer	166
Db	910	GATGGCAGAGTTATTTCTTCTGTTGGTAAAGTACGGGAAGGATGATGGGTAGC	969
Qy	167	IleSerArgAspPheValIleSerArgThrProThrGlnValAlaSerHisAlaGln	186
Db	970	ATTCTCGCACTTGTAGTAAACAAACACGAGCCAGATGCGACGATCTCAAAAG	1029
Qy	187	TyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerSerIleHisAspIle	206
Db	1030	TATTTCAATCGCTAAATTCATGAAACAAAGCAGAGCGATCAAGCATTACAGACATC	1089
Qy	207	ThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGlyHis	226
Db	1090	ACTAGTGTGGCAACGACAGAT---GTCTCAACGCCACAAAGA---CAATCACTGTGTCAG	1143
Qy	227	GlnAlaThrGlyAsn-----ProAlaAlaAlaAla	236
Db	1144	AACACACACAAATACACACACACACACACACACAGTTCTCCGCTGTTGCTGCA	1203
Qy	237	LeuGlyProGlyMetLyHisIleHisIleHisIleProGlyGlyAlaProProPro	256
Db	1204	GGAGGAACAATACATCCAG---CAAGCGCTCTCCAGCACCACCTGGA	1251
Qy	257	MetProMetTyrSerAlaAlaProMetGlyHisProValAlaGlyHisMetValProAla	276
Db	1252	CCTCTATGTATGGAACACCCGCCCTAGTGTGTCACCA-----	1287
Qy	277	AlaValGlyThrProValValPhePro---ProGlyHisAlaProTyrValVal	293
Db	1288	GCAGTGGACACCGATGAACCTCCAGCTCAGCTCAGATGGCTTATGAGATTCGCG	1347
Qy	294	---ProVal---GlyTyrProAlaProProProAlaLyMetHis	305
Db	1348	GCTCAGTCCCTCGCTCAGTGTCTGTGTGTCAGCAATGAAC	1389
RESULT 14			
AK068138	1988 bp	mRNA	linear
LOCUS			PLN 24-JUL-2003
DEFINITION	Oryza sativa (japonica cultivar-group)	cDNA clone:J013135D01, full	
ACCESSION	AK068138		
VERSION	AK068138.1	GI:32978156	
KEYWORDS	FLU CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE			
AUTHORS			

TITLE	
JOURNAL	Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
MEDLINE	Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Otsu, N., Oka, Y.,
PUBMED	Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
REFERENCE	Yoshino, M. and Hayashizaki, Y.
AUTHORS	Collection, mapping, and annotation of over 28,000 cDNA clones from
	Japonica rice
	Science 301 (5631), 376-379 (2003)
	2 (bases 1 to 1988)
	12869764
	2 (bases 1 to 1988)
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
	Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
	Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirao, T.,
	Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
	Imotani, K., Ishibiki, T., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
	Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
	Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
	Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
	Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
	Maeda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
	Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
	Namiki, T., Narikawa, R., Nikura, J., Niishi, K., Nomura, K.,
	Nunasaki, R., Ohneda, E., Ono, M., Ohtsuki, K., Oka, M., Ooka, H.,
	Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
	Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
	Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
	Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
	Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
	Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W.,
	Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
	Yoshimura, A.
TITLE	Direct Submission
JOURNAL	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
	Agricultural Sciences, Department of Molecular Genetics, Head of
	Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
	305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp).
	Tel:81-29-838-7007, Fax:81-29-838-7007
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica
	rice.
	URL : http://cdna01.dna.affrc.go.jp/cDNA/
	NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
	Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
	Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Shishiki, T.,
	Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
	Yamamoto, M.
	FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,
	Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
	Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Maeda, H., Miura, J.,
	Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S.,
	Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
	Yoshimura, A., Matsubara, K. and Murakami, K.
	Genome Exploration Research Group in Riken Genomic Sciences Center
	and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
	Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
	Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
	Hirao, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Iishi, Y.,
	Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
	Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
	Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Miura, T.,
	Nakamura, M., Niishi, K., Nomura, K., Nunasaki, R., Ono, M., Otsu, N.,
	Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
	Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
	Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
	Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
	Yasunishi, A. and Hayashizaki, Y.
FEATURES	Location/Qualifiers
source	1..1988
	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="mRNA"
	/cultiyar="Nipponbare"
	/db_xref="taxon:39947"
	/clone="J013135D01"
ORIGIN	



```

Db      194  TGAAGTCTGCTACCAAGCAAGCTATTGAGCGCGCCCTTCTGTC----- 238
Qy      38   ProAlaAspGlyGlyAlaProAspAsp-----AspTrpPheAlaAlaLeuAlaAlaSer 55
Db      239  -----GTCCCGAGAGATCTCCCGGACCGGTGGAGAGAGATGCGCGACG 283
Qy      56   ValProGlyAlaArgSerAlaGluGluValArgArgHisTyrGluAlaLeuValGluAsp 75
Db      284  GTTCCCGGC---AAGTCGCGGTGAGAGTAAGGGAACACTACGAGGCCCTGTCACGAC 340
Qy      76   ValAlaAlaIleAspAlaGlyArgValProLeuProArgTyrAlaGlyGluGluSerAla 95
Db      341  GTCTTCGAGATCGACTCCGCGGAGTCCGAGTCCGAGTTACGTGATGACTCCGCTGCGC 400
Qy      96   AlaProProAspGlyAlaGlyAlaAlaAlaAlaSerLysAspGlyGlyHisArgArg 115
Db      401  ATGCGGCCC-----AGTGGCGGTGTGTGAAATTCTACGTGGACACGCTAATCAGATC 454
Qy      116  AspGluArgLysGlyGlyGlyGlyTyrAspGlyGlyLysSerCysSerLysAlaGlu 135
Db      455  TCA-----TTCCGTTCCAGCTTAACAACAAGCGCAG 487
Qy      136  GlnGluArgArgLysGlyIleProTrpThrGluGluGluHisArgLeuPheLeuGlu 155
Db      488  AACGAAAGGAAGAGGCGACGCTTGACGACGAAACACAGCGCTGTTCTCATGCG 547
Qy      156  LeuAspLysPheGlyLysGlyAspTrpArgSerLieserArgAsnPheValIleSerArg 175
Db      548  CTAGTAGTGTGGAAAGGAGATTGAGAAAGCATTTCAAGAAATGTTGTGTGCAAGA 607
Qy      176  ThrProThrGluValAlaSerHisAlaGlnLysTyrPheIleArgLeuAsnSerMetAsn 195
Db      608  ACACCAACTCAAGTAGTACGATGCCAAGAAATCTTCTTCGCCAGAAATTCGCTGAAG 667
Qy      196  ArgAspArgArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGluVal 215
Db      668  AAAGAGAGAGAGATCAAGCATTCATGATTAACGACTGTGACAGT----- 715
Qy      216  AlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGlyAsnProAlaAlaAla 235
Db      716  -----AACTCAGCGCCTATG-----CCCATGTGATCAG 742
Qy      236  AlaLeuGlyProProGlyMetLysHisHisHisHisHisProGlyGlyAlaProPro 255
Db      743  ACTTGGGTTCCTCT-----CCGGTGGTTCACCTCAG 775
Qy      256  ProMetProMetTyrSerAlaAlaProMet 265
Db      776  CAGTCACAGCAATACCATCAAGCAATG 805

```

Search completed: July 30, 2005, 00:26:00  
 Job time : 5168 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - nucleic search, using frame\_plus\_p2n model

Run on: July 29, 2005, 19:31:08 / Search time 614 Seconds  
(without alignments)  
2950.229 Million cell updates/sec

Title: US-10-630-636-7  
Perfect score: 1641  
Sequence: 1 MTSQAATTTTAAATAAATWTR.....GHAPVVPVGYPPAPAKMHQ 306

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh  
-O=/cgn2.1/USPFO\_epoc1/US10630636/runat\_27072005\_123226\_12158/app.query.fasta\_1.455  
-DB=N Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFCI=0  
-LOOPEXT=0 -UNITIS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10630636@cgn.1.1.644@runat\_27072005\_123226\_12158 -NCPU=6 -ICPU=3  
-NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-DEV TIMEOUT=150 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:  
1: geneseqn1980s:  
2: geneseqn1990s:  
3: geneseqn2000s:  
4: geneseqn2001as:  
5: geneseqn2001bs:  
6: geneseqn2002as:  
7: geneseqn2002bs:  
8: geneseqn2003as:  
9: geneseqn2003bs:  
10: geneseqn2003cs:  
11: geneseqn2003ds:  
12: geneseqn2004as:  
13: geneseqn2004bs:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1641	100.0	918	12	AD000780 Rice Myb
2	1641	100.0	1330	12	AD000777 Rice Myb
3	678.5	41.3	1161	3	AC34213 Arabidops
4	624.5	38.1	1432	10	AD04636 Thalecres
5	624.5	38.1	1432	10	AD55829 Thalecres

6	624.5	38.1	1432	12	AD161458	Adi61458 cDNA enco
7	624.5	38.1	1432	12	AD002450	Ado002450 Thalecres
8	457.5	27.9	867	6	ABK65275	ABK65275 Arabidops
9	457.5	27.9	867	10	AD030621	Ad030621 Plant yie
10	457.5	27.9	867	10	AD031552	Ad031552 Plant yie
11	457.5	27.9	867	12	AD144216	Ad144216 Plant tra
12	456	27.8	668	6	ABK65273	ABK65273 Arabidops
13	456	27.8	668	10	AD030307	Ad030307 Plant yie
14	456	27.8	668	12	AD144214	Ad144214 Plant tra
15	456	27.8	668	12	AD144214	Ad144214 Plant tra
16	455.5	27.8	608	13	ACN45747	ACN45747 Cotton pr
17	445	27.1	891	3	AC39215	AC39215 Arabidops
18	428.5	26.1	866	6	ABK65370	ABK65370 Arabidops
19	428.5	26.1	866	10	AD030950	Ad030950 Plant yie
20	428.5	26.1	866	10	AD037148	Ad037148 Plant yie
21	428.5	26.1	866	12	AD141988	Ad141988 Plant tra
22	428.5	26.1	866	12	AD003516	Ad003516 Thalecres
23	428.5	26.1	866	3	AC47819	AC47819 Arabidops
24	423	25.8	804	10	AD029992	Ad029992 Plant yie
25	423	25.8	804	12	AD144406	Ad144406 Plant tra
26	420.5	25.6	896	6	ABK65271	ABK65271 Arabidops
27	420.5	25.6	896	10	AD046650	Ad046650 Thalecres
28	420.5	25.6	896	10	AD030965	Ad030965 Plant yie
29	420.5	25.6	896	12	AD141842	Ad141842 Plant tra
30	420.5	25.6	896	12	AD002514	Ad002514 Thalecres
31	419.5	25.6	512	12	AD033595	Ad033595 Cotton ex
32	412.5	25.1	574	3	AA056812	AA056812 Eucalyptu
33	390	23.8	461	12	AD093427	Ad093427 Cotton ex
34	386.5	23.6	335	3	AA056762	AA056762 Eucalyptu
35	381.5	23.2	576	13	ACN49160	ACN49160 Cotton pr
36	368.5	22.5	392	12	AD091116	Ad091116 Cotton ex
37	367	22.4	580	13	ACN48912	ACN48912 Cotton pr
38	363.5	22.2	559	13	ACN49230	ACN49230 Cotton pr
39	351.5	21.4	1287	12	AD000779	Ad000779 Rice Myb
40	344	21.0	954	12	AD000782	Ad000782 Rice Myb
41	336.5	20.5	332	3	AA057102	AA057102 Pinus rad
42	335.5	20.4	1354	3	AA046782	AA046782 Arabidops
43	335.5	20.4	1386	10	AD031814	Ad031814 DNA encod
44	335.5	20.4	1386	12	AD002038	Ad002038 Thalecres
45	333.5	20.3	349	3	AA056698	AA056698 Eucalyptu

## ALIGNMENTS

RESULT 1	
ID	AD000780 standard; cDNA, 918 BP.
AC	AD000780;
XX	
DT	26-AUG-2004 (first entry)
XX	
DE	Rice Myb protein OSMYBS1 CDS cDNA.
XX	
KW	ss; gene; OSMYBS1; rice; plant; transgenic; Myb; transcription factor.
XX	
OS	Oryza sativa.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..918
FT	/*tag= a
FT	/partial
FT	/product= "OSMYBS1"
FT	/note= "No stop codon given"
XX	
PN	US2004107456-A1.
XX	
PD	03-JUN-2004.
XX	
PF	30-JUL-2003; 2003US-00630636.
XX	
PR	31-JUL-2002; 2002US-03999999.
XX	







PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 16-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139482P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140685P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145152P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148655P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149446P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158212P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159684P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.









XX	(ZHAN/	ZHANG J.	
PA	(FROM/	FROMM M.E.	
PA	(HEAR/	HEARD J.E.	
PA	(RIEC/	RIECHMANN J.L.	
PA	(ADAM/	ADAM L.J.	
PA	(BROU/	BROWN P.E.	
PA	(PINE/	PINEDA O.	
PA	(REUB/	REUBER T.L.	
PA	(KEDD/	KEDDIE J.S.	
PA	(YUGG/	YU G.	
PA	(JTAN/	JTANG C.	
PA	(SAMA/	SAMAH R.S.	
PA	(PILG/	PIELGRIM M.L.	
PA	(CREE/	CREELMAN R.A.	
PA	(DUBE/	DUBELL A.N.	
PA	(RATC/	RATCLIFFE O.	
PA	(KUMI/	KUMIMOTO R.	
PA	(SHER/	SHERMAN B.K.	
XX			
P1	Zhang J.	Fromm ME,	Heard JE, Riechmann JL, Adam LJ, Brown PE;
P1	Pineda O,	Reuber TL,	Keddle JS, Yu G, Jiang C, Samaha RS;
P1	Pilgrim ML,	Creelman RA,	Dubell AN, Ratcliffe O, Kumimoto R;
P1	Sherman BK,		
XX			
DR	WPI:	2004-225755/21.	
XX	P-PSDB;	AD002451.	
PT			
PT	New transgenic plant,	useful in developing phenotypes with altered or	
PS	improved characteristics or traits.		
XX			
XX	Claim 1;	SEQ ID NO 863;	213pp; English.
CC			
CC	The invention relates to a transgenic plant comprises a recombinant		
CC	polynucleotide having a polynucleotide sequence or its complementary		
CC	sequence comprising a sequence encoding a polypeptide, that initiates		
CC	transcription (i.e. a transcription factor) from Arabidopsis, Soybean,		
CC	Rice, Rape or Corn, comprising any of the sequences appearing as AD0015888		
CC	-AD003327 or AD003530-AD003559. Also included are using a transgenic		
CC	plant to grow a progeny plant, an expression cassette (comprising a		
CC	constitutive, inducible or tissue-specific promoter and a recombinant		
CC	polynucleotide described above), a host cell comprising the expression		
CC	cassette, producing a modified plant having a modified trait, identifying		
CC	a factor that is modulated by or interacts with a polypeptide encoded by		
CC	the polynucleotide sequence and identifying at least one downstream		
CC	polynucleotide sequence that is subject to a regulatory effect of any of		
CC	the polypeptides encoded by the polynucleotide described above. The		
CC	transgenic plant is useful for producing a plant that has an altered		
CC	trait e.g. an enhanced tolerance to abiotic stress (increased tolerance		
CC	to chilling, germination in cold conditions,freeing tolerance, tolerance		
CC	to heat, tolerance to drought, tolerance to osmotic stress, tolerance to		
CC	salt, tolerance to phosphate limitation, tolerance to potassium		
CC	limitation, decreased sensitivity to nitrogen limitation), altered		
CC	hormone sensitivity, reduced sensitivity to abscisic acid, an altered		
CC	response to ethylene, disease resistance, altered susceptibility to		
CC	Borerytis, altered susceptibility to Fusarium, altered susceptibility to		
CC	Erysiphe, altered susceptibility to Pseudomonas syringae, altered		
CC	susceptibility to Sclerotinia, altered sugar sensing, improved seed		
CC	germination and seedling vigor, early flowering, late flowering, extended		
CC	period of flowering, an inflorescence architectural change, a change in		
CC	stem bifurcations, a lack of a shoot meristem, reduced meristem cell		
CC	division, altered phylloclaxy, altered branching pattern, reduced		
CC	apical dominance, reduced trichome density, ectopic trichome development,		
CC	altered trichome development, altered stem morphology, increased root		
CC	growth, increased root hairs, altered seed development, altered cell		
CC	proliferation/cell differentiation, premature senescence, delayed		
CC	senescence, lethality, increased necrosis, an increase in seedling or		
CC	plant size, decreased plant size, a change in leaf morphology, increased		
CC	altered leaf development, increased leaf size and mass, glossy leaves,		
CC	leaf cell expansion, change in seed morphology, altered seed coloration,		
CC	increased seed size, decreased seed size, altered seed shape, change in		
CC	leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid		
CC	content, increased leaf insoluble sugars, decreased leaf insoluble		

[illegible]











PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,  
PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
XX WPI: 2002-292022/33.  
DR P-PSDB: AAU93087.  
XX  
PT An isolated or recombinant polynucleotide used to produce a transgenic  
PT plant.  
XX  
PS Claim 4; Page 521-522; 941pp; English.  
XX  
CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
CC encoding an Arabidopsis thaliana transcription factor, their variants,  
CC complements, fragments, or related polynucleotide with 31% to 95%  
CC sequence identity, where the plant possesses an altered trait as compared  
CC to a wild-type or reference plant, or the plant exhibits an altered  
CC phenotype as compared to a wild-type or reference plant, or the plant  
CC exhibits ectopic expression or altered expression of one or more genes  
CC associated with a plant trait as compared to a wild plant. Also included  
CC are a transgenic plant comprising the polynucleotides, a computer  
CC readable medium having stored sequence information, and identifying a  
CC homologous sequence from a database comprising a plurality of known plant  
CC sequences comprising inputting sequence information selected from one of  
CC 466 fully defined sequences given in the specification. The isolated or  
CC recombinant polynucleotide is used for producing a plant having a  
CC modified trait, the method comprising selecting a polynucleotide that  
CC encodes a polypeptide or an antisense nucleic acid, inserting the  
CC polynucleotide or antisense nucleic acid into an expression vector,  
CC introducing the vector into a plant or a cell of a plant to overexpress  
CC the polypeptide or antisense nucleic acid, thereby producing a modified  
CC plant, and selecting for a modified trait (e.g. increased production of  
CC agriculturally useful proteins or metabolic chemicals, pest tolerance,  
CC environmental stress response (e.g. drought), microbial disease  
CC resistance, herbicide resistance, seed and fruit yield, growth rate, leaf  
CC and flower senescence and many other traits listed in the specification).  
CC The present sequence is one of the 232 polynucleotides encoding an A.  
CC thaliana transcription factor  
CC  
XX  
XX Sequence 668 BP; 170 A; 119 C; 204 G; 175 T; 0 U; 0 Other;  
XX  
XX  
Alignment Scores:  
Pred. No.: 6 95e-18 Length: 668  
Score: 456.00 Matches: 101  
Percent Similarity: 58.77% Conservative: 33  
Best Local Similarity: 44.30% Mismatches: 52  
Query Match: 27.79% Indels: 42  
DB: 6 Gaps: 7  
US-10-630-636-7 (1-306) x ABK65273 (1-668)  
QY 9 ThrtThrThraAlaAlaAlaAlaAlaAlaTrPhrXrGluAerApySaAlaPhaGluasn 28  
Db 7 TCATCAAGCATTAATGGGTCAAGTCAGTCGACGAGTTCGAGATTAAGATTTTGAGCAA 66  
QY 29 AlaleuAlaAlaCysAlaAlaProProRoAlaAspGlyAlaProAerAspAspTtr 48  
Db 67 GCTTGGTCTTTTCTCTGAA-----GGATCTCCATATCG----- 102  
QY 49 PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluGluValArgXHis 68  
Db 103 TGGGAGCAATCGCTGATCAGCTT-----CATAAATCGCTGGTGAAGTTAAGGAGCAT 156  
QY 69 TyrGluAlaLeuValGluAspValAlaAlaAlaLeaAspAlaGlyArgValProLeuProArg 88  
Db 157 TACGAGGCTTGGTTCATGATGTTTTCAGATGATGATTCGTGCAAGTTGATGATCGAT 216  
QY 89 TyrAlaGlyGluGluSerAlaAlaProPro-----AspGlyAlaGlyAlaAlaAla 106  
Db 217 TACATGAGTACGTCGCGCTGCGCGCGCGCTGGTGGATTCGCGTGCAGATCTCTTT 276  
QY 107 AlaSerLysAspArgGlyGluHisArgAspRluArgLysGlyGlyGlyGlyGlyTyrAsp 126  
Db 277 GGGTCTAAACATGGC----- 291

QY 127 GlyGlyLysSerCysSerLysAlaGluGluArgArgGlyGlyIleProTrrThrGlu 146  
Db 292 -----GAGAGTGAACGCAAAAGAGCAACTCCTTGACACAG 327  
QY 147 GluGluHisArgLeuPheLeuGlyLeuAspLysPheGlyLysGlyAspTrrPArgSer 166  
Db 328 AACCAACAAATTTGTTTCTGATCGATTAAAGAGATATGTTAGGAGATGGAGAGT 387  
QY 167 ILeSerArgAsnPheValIleSerArgThrProThrGluValAlaSerHisAlaGlnLys 186  
Db 388 ATCTCGAGAAACGTTGTGTGAGACGACCGACCGAATCGCGATCGCTCAGAG 447  
QY 187 TyrPheIleArgLeuAsnSerMetAsnArgAspArgArgArgSerSerIleHisAspIle 206  
Db 448 TATTTCTGAGACGAACCTCGTGAAGAGAGGAAAGTTCGAGCATTCATGATATA 507  
QY 207 ThSerValThrAla-----GlyAspGlnVal-----AlaAlaGlnGln 219  
Db 508 ACTACGGTTATGCTACTTGTGCTATGCTGGTCTTAACTGAACCTGAGCTGCGCAACAC 567  
QY 220 GlyAlaProIleThrGlyHisGln 227  
Db 568 GCGAGTCTTTCAGGCGCCGACG 591  
RESULT 13  
ADD30307  
ID ADD30307 standard; cDNA; 668 BP.  
XX  
XX ADD30307;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Plant yield-related polynucleotide clone G1636.  
XX  
XX de; transcription factor; transgenic plant; growth rate; senescence;  
XX seed germination rate; plant vigor; seedling vigor.  
XX  
XX Arabidopsis thaliana.  
XX  
XX W02003013227-A2.  
XX  
XX 20-FEB-2003.  
XX  
XX 09-AUG-2002; 2002WO-US025805.  
XX  
XX 09-AUG-2001; 2001US-0310847P.  
XX 19-NOV-2001; 2001US-0336049P.  
XX 11-DEC-2001; 2001US-0338692P.  
XX 14-JUN-2002; 2002US-00171468.  
XX  
XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
XX  
XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AJ, Heard JB;  
XX Pilgrim M, Jiang C, Reuber JL, Creelman RA, Pineda O, Yu G;  
XX Brown PE;  
XX  
XX WPI: 2003-248221/24.  
XX P-PSDB: ADD30308.  
XX  
XX New plant transcription factor polynucleotides and polypeptides, useful  
XX in producing transgenic plants with commercially valuable properties,  
XX PT such as an alteration in a plant growth characteristic, e.g. growth rate  
XX or apomixis.  
XX  
XX Disclosure; SEQ ID NO 336; 454pp; English.  
XX  
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA  
XX sequences and their encoded proteins which are especially transcription  
XX factor related cDNA's and proteins. The isolated or recombinant plant  
XX transcription factor polynucleotides and polypeptides are useful in  
XX producing transgenic plants with commercially valuable properties, i.e.  
XX modified or altered desirable traits as compared to a reference plant.







```
QY 107 AlaSerIysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGlyGlyTyrAsp 126
    ||| |||
    ||| |||
Db 390 GGGCTTAACATGGC----- 404
QY 127 GlyGlyLysSerCysSerLysAlaGluGluArgArgLysGlyLeuProIleThrGlu 146
    ||| |||
    ||| |||
Db 405 -----GAGAGTGAAACGCAAAAGAGAACTCCTTGACACAG 440
QY 147 GluGluHisArgLeuPheLeuGlyLeuAspLysPheGlyLysGlyAspTyrArgSer 166
    ||| |||
    ||| |||
Db 441 AACGAAACAATAATTGTTCTGATCGATTAAAGATATATGTTAAGGAGATTGAGAGAGT 500
QY 167 IleSerArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLys 186
    ||| |||
    ||| |||
Db 501 ATCTCGAGAAACGTTGTGTGACGAGACACGCAAGTCGCGAGTCACGTCAGAAAG 560
QY 187 TyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerIleHisAspIle 206
    ||| |||
    ||| |||
Db 561 TATTTTCTGACACAGAACTCGGTGAGAAGAGAGAAAGGTGAGCATTCATGATATA 620
QY 207 ThrSerValThrAla-----GlyAspGlnVal-----AlaAlaGlnGln 219
    ||| |||
    ||| |||
Db 621 ACTACGGTTGATGCTTGTGCTATGCGCTGCTTAACATGACTGACTGCGCCACAC 680
QY 220 GlyAlaProIleThrGlyHisGln 227
    ||| |||
    ||| |||
Db 681 GGGAGTCCTGTTTCAGGCGCCGACG 704
```

Search completed: July 29, 2005, 22:59:56  
Job time : 625 secs





```
QY 34 AlAlaIaProProAlaAspGlyAlaIaProAspAspPheAlaIaLeuAla 53
    |||
Db 377 TTCAGAGTTTCCCGC-----GATTGTTGGAGAAAGTTAGCG 342
QY 54 AlAspValProGlyAlaArgSerAlaGluValaArgAghIserGluAlaLeuVal 73
    |||
Db 341 TCCGACGTGGCCGGC---AAAACCTTAGAGAGATTAAAGCTCCACCTACGAGACCTGGTG 285
QY 74 GluAspValAlaAlaIaAspAlaGlyAArgValProLeuProArgGlyAlaGluGlu 93
    |||
Db 284 GAAGATGTCATAGATAGAGAGCGTGGCGGCTGCTGCGCGTCTAC----- 237
QY 94 SerAlaIaProProAspGlyAlaGlyAlaAlaIaAlaAspSerIAspGlyHis 113
    |||
Db 236 AGTTCCGCTTGTCTCGATGGC-----TCCGCGCAAGTACTTAGGGGAG 195
QY 114 ArgArgAspGluArgLysGlyGlyGlyTyrAsp-----GlyGlyLysSer 130
    |||
Db 194 GAGGAGAGCAAGAAAGGGGGCGGCGGTAGCGCGCAAGTCMAACGAGGAGGAAAGGC 135
QY 131 CysSerIAspAlaGluGluGluArgArgLysGlyIleProTrrPthrGluGluHisArg 150
    |||
Db 134 ---TCGAGGTCCGATCAGAGAGAGAGAGGAGATTGCTTGACCGAGAAAGACAGG 78
QY 151 LeuPheLeuGluGlyLeuAspLysPheGlyLysAspTrrPargSerIleSerArgAsn 170
    |||
Db 77 TTAATTTCTTAGGCTTAGATTAATATAGGAAAGCGATGGCGAATATTTCCCGAAGC 18
QY 171 PheValIleSerArg 175
    |||
Db 17 TTGTGCTCACAAAG 3
```

```
RESULT 2
US-09-640-211A-1484
; Sequence 1484, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1484
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(335)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1484
```

```
Alignment Scores:
Pred. No.: 2.18e-23 Length: 335
Score: 386.50 Matches: 78
Percent Similarity: 80.91% Conservative: 11
Best Local Similarity: 70.91% Mismatches: 8
Query Match: 23.55% Indels: 13
DB: 4 Gaps: 3
```

US-10-630-636-7 (1-306) x US-09-640-211A-1484 (1-335)

```
QY 116 AspGluArgLysGlyGlyGlyGly-----TyrAspGlyGlyLysSerCys 131
    |||
Db 7 GAGAGCAGAGAAAGGGGGCGGCGGTAGCGCGCAAGTCMAACGAGGAGGAAAGGC--- 63
QY 132 SerIAspAlaGluGluGluArgArgLysGlyIleProTrrPthrGluGluHis----- 149
```

```
Db 64 TCGAGGTCCGATCAGAGAGAAAGGGGATTGCTTGACCGACGAGAAACACAGCTGG 123
QY 150 -----ArgLeuPheLeuGlyLeuAspLysPheGlyLysArg 163
    |||
Db 124 GGTCTTCTTCTTGCAATCGGTTATTCTTCTTAGGCTTAGTAAATATGGAAAGGGGAT 163
QY 164 TrrPargSerIleSerArgAsnPheValIleSerArgThrProThrGluValAlaSerHis 183
    |||
Db 184 TGGCGAAGTATTTCCCGAACTTTGTTCGNCAAGGACCGCTACGCAAGTTGGAGGCAT 243
QY 184 AlArgLysPheThrPheIleAspLeuAsnSerMetAsnArgAspArgArgArgSerIle 203
    |||
Db 244 GCACAGAGATATTTTTCGTGAGCTCTGTTAACCAAGATAGAGCGCATCTAGCATT 303
QY 204 HisAspIleThrSerValThrAlaGlyAsp 213
    |||
Db 304 CATGATTTCACTACCGTAGGCACTGAGAGAC 333
```

```
RESULT 3
US-09-640-211A-1824
; Sequence 1824, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1824
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1824
```

```
Alignment Scores:
Pred. No.: 2.68e-19 Length: 332
Score: 336.50 Matches: 72
Percent Similarity: 71.07% Conservative: 14
Best Local Similarity: 59.50% Mismatches: 23
Query Match: 20.51% Indels: 13
DB: 4 Gaps: 3
```

US-10-630-636-7 (1-306) x US-09-640-211A-1824 (1-332)

```
QY 64 GluValaArgAghIserGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArg 83
    |||
Db 4 GAGGTAGAGAGCATTTTCGAGCTTCTTTGAGAGATGTGCTGTGATTGAGTCTGCGCGG 63
QY 84 ValProLeuProArgGlyAlaGluGluGluSerAlaIaProProAspGlyAlaGlyAla 103
    |||
Db 64 GTTGCTTGCTGCTGCTATTTCT---GAAATTCGATACACCGCCCGCAATGATGTACA- 119
QY 104 AlAlaIaAlaAspSerIAspGlyHisArgArgAspGluArgLysGlyGlyGly 123
    |||
Db 120 TCAAGTTGGCGGACT-----CACAAACA-----GCAGGC 149
QY 124 GlyTrrAspGlyGlyLys-SerCysSerIAspAlaGluGluGluArgArgLysGlyIlePr 143
    |||
Db 150 GGTTCCTGTGAAGCTCTCTCGCCCAAGCATCGAAGCAGACGCAAAAAGGCGCGCC 209
QY 143 OTrrPthrGluGluGluHisArgLeuPheLeuGluGlyLeuAspLysPheGlyLysArg 163
    |||
Db 210 CTGAGCTGAAGAGACAGACACTCTTCTTAGGAGGATTAATTAATATATGGCAAAAGTGA 269
QY 163 PrrPargSerIleSerArgAsnPheValIleSerArgThrProThrGluValAlaSerHis 183
    |||
```





Db 450 GGTAAAGACGCTGGCGAGGATACCTGCTGATCTTGTGACTACAAAGACTCTCTACTCA 509  
Qy ValAlaSerHisAlaGlnLeuTyrPheIleArgLeuAsnSerMetAsnArgAspArg 199  
Db 510 GTGGACACCTGCTCCGAGAGTATTATTCGGCAGAGTAACTGCGCCGAAAGAGAGG 569  
Qy 200 ArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGln 219  
Db 570 CGCTCAGCTCTTTGACATG-----GCTCCAGATATGCGTACTGCTGACCA 617  
Qy 220 GlyAla-----ProIleThrGly---HisGlnAlaThrGly 230  
Db 618 CCGTCACATCCAGAAAGAAATCTTCTGCTCTTGTGACAGACTTAAGCATGATCTAC 677  
Qy 221 AsnProAlaAlaAlaLeuGly-----ProProGlyMetLeuHisHis 245  
Db 678 TCAACAACTTCAACCGATGTGGACTCGATTTGGAAGAACCCCTATGGAGACCTCGAC 737  
Qy 246 HisHisHisHisProGlyGlyAlaProProPromeTyrSerAlaAlaProMet 265  
Db 738 CCAGAAACATCGAAGGGGGCGGTATGTGGATGGAATCAATTGATCAAGTACTCTT 797  
Qy 266 GlyHisProValAlaGlyHisMetValProAlaAlaValGlyThrProValAlaPro 285  
Db 798 -----GTA 800  
Qy 286 ProGlyHisAlaProTyrValValProValGlyTyrProAlaProProAlaLeuMet 304  
Db 801 CCTGTTACTTCCACTATTACCACTACCTTTCCATGTGGCGCCCAACATG 857  
RESULT 9  
US-09-640-211A-232  
; Sequence 232, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: Modification of Gene Transcription  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 232  
; LENGTH: 1358  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-09-640-211A-232  
Alignment Scores:  
Pred. No.: 3,19e-17 Length: 1358  
Score: 321.00 Matches: 109  
Percent Similarity: 40.00% Conservatve: 25  
Best Local Similarity: 32.54% Mismatches: 94  
Query Match: 19.56% Indels: 108  
Gaps: 15  
US-10-630-636-7 (1-306) x US-09-640-211A-232 (1-1358)  
Qy 34 AlaAlaPro---ProProAlaAsp-----GlyGlyAlaProAspAspAspTyrPheAla 50  
Db 118 GCGAGTCCCGACCCCGCGCGCATGACGCGCGATGCTCCCACT-----GCT 165  
Qy 51 AlaAlaAlaAlaSerValProGly---AlaArgSerAlaGlnGlnValAlaArgHisTyr 69  
Db 166 GCAACAGAGGCGCACACTCGACCTGCGCCCTCGCG----- 204  
Qy 70 GluAlaLeuValGluAspValAlaAlaIleAspAlaGly-----ArgValProLeuPro 87  
Db 205 -----GCGGCGGCGGGAACGCGGCGGCGCGCGCGCGCCCTCT 246

Qy 88 ArgTyrAlaGlyGlnGluSerAlaAlaProProAspGlyAlaGlyAlaAlaAlaAla 107  
Db 247 -----CCTCCTCCCTCCCACTCTCTCTCT-GGCGCGCGCGCGCGCGCGCGCC 296  
Qy 108 SerIleAspGlyGly----- 112  
Db 297 TCGGCTCCGCGCGCGGCGGTAAGCTGTTCCGGGTTAGATTAAACGACGGGTGATCATG 356  
Qy 113 -----HisArgHisArgGluArgGly 120  
Db 357 AAGAAAGCGCCAGCGGTGGGTGCTCTCGCCGCCCACTACCTCTCTCTCCGCC 416  
Qy 121 GlyGlyGly-----GlyTyrAspGly----- 127  
Db 417 GCGGATCCCGGAACCCCGGCTGTCCTCCCATGACGGAGACGACGCTTACTCTCGAC 476  
Qy 128 -----GlyIleSerCysSerIleAlaGlnGlnGluArgArgGlyTyrPro 143  
Db 477 GATCCCGCGCGCGCTCCGCTCCCATGCGCGCGCTCGAGAGGAAAGTAAACCA 536  
Qy 144 TrpThrGlnGlnGlnHisArgLeuPheLeuGlnGlyLeuAspIlePheGlyIleAsp 163  
Db 537 TGAAGGAGAGAGACATCGAAGGTTTAAATGGTCTCCAGAAATTGGTAAAGAGAC 596  
Qy 164 TrpArgSerIleSerArgAsnPheValIleSerArgThrProThrGlnValAlaSerHis 183  
Db 597 TGGCGAGGATAGCTCGGACTTGTACTGACACTCAAGACTCTTCACTCAAGTGGCAAGCAT 656  
Qy 184 AlaGlnIleTyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerSerIle 203  
Db 657 GCCGAGAGTATTTATCTCCGACAGATAGCTGGCCGAAAGAGAGCGCTCCAGCTT 716  
Qy 204 HisAspIleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAla----- 221  
Db 717 TTGACATG-----GCTCCAGATATGCTACTGTGACCAACCTCTCATATCCA 764  
Qy 222 -----ProIleThrGly---HisGlnAlaThrGlyAsnProAlaAla 234  
Db 765 GAAGAAACATTTGCTGCTCTTGTGCTGACACTTAACGATGATCTACTCAACAACCTTCA 824  
Qy 235 AlaAlaLeuGly-----ProProGlyMetLeuHisHisHisHisHisHis 249  
Db 825 ACCAGTATGGACATCGATTTGGAAGACGCTATGAGACCTCGCACCCGAAACATCT 884  
Qy 250 ProGlyGlyAlaProProPromeTyrSerAlaAlaProMetGlyHisProVal 269  
Db 885 GAAGGGGCGGTGATGTTGCGATGGAATCAATTGATCAAGTACTCTT----- 932  
Qy 270 AlaGlyHisMetValProAlaAlaValGlyThrProValAlaPheProProGlyHisAla 289  
Db 933 -----GTACCTGTACTTTC 947  
Qy 290 ProTyrValValProValGlyTyrProAlaProProAlaLeuMet 304  
Db 948 CCAATCATTTTACCACTACCTTTCATGCGCGCGCAACATG 992  
RESULT 10  
US-09-640-211A-2038  
; Sequence 2038, Application US/09640211A  
; Patent No. 6833446  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: Modification of Gene Transcription  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2368  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 2038
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-2038

Alignment Scores:
Pred. No.: 1,51e-16 Length: 1055
Score: 311.00 Matches: 70
Percent Similarity: 52.81% Conservative: 24
Best Local Similarity: 39.33% Mismatches: 42
Query Match: 18.95% Indels: 42
DB: 4 Gaps: 5

US-10-630-636-7 (1-306) x US-09-640-211A-2038 (1-1055)
QY 71 AAlaevAlaGluAspValAlaAlaAlaAspAlaGlyArgValProLeuProArg----- 88
Db GCTCTGCTCTTCAATTGGGTTCTGTGACCGAGCTCCATCAACAATGTCAAAGATTG 358
QY 89 -----TyrAlaGlyGluGluSerAlaAlaProProAspGly 100
Db TTGATGATGTGTTCACACTGTGGCAGACAGTGGCAGACAGCTCCAGACCTGTCTGATAGA 418
QY 101 Ala-----GlyAlaAlaAlaAlaAlaSerLysAspGly-----GlyHis 113
Db GGATCTGTCAAAATGTTTGGGTCAGGCTCATTTCTACAGACGATGGCATGGCTGCATG 478
QY 114 ArgArgAspGluArgLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 125
Db AGAAGAGACCTCAGATATGGGCAATCTCGTCATTAACGCTTACATTAATGCAATCAC 538
QY 126 ---AspGlyGlyLysSerCysSerLysAlaGluGln----- 136
Db TGTCTGGACAAACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 538
QY 137 -----GluArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 148
Db GTTCATTTCTCCAGCAATGCACGGCAGAGAAAAAGCCGCCATGCTGTCGAGAGAG 658
QY 149 HisArgLeuPheLeuLeuGlyLeuAspLysPheGlyAspLysPheGlyAspLysPhe 168
Db CACGAGATGTTCTTGTATGACCTGAAAAGCTTGGAAGGTGCTGGAAGGATATCC 718
QY 169 ArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLysTyrPhe 188
Db AGGAATTTTGTGACGACGAAACACCAACCAAGTACGACGACCAAGTATTTT 778
QY 189 IleArgLeuAsnSerMetAsnArgAspArgArgArgSerIleHisAspIle 206
Db 779 CTAAAGCAGACCAATCTTAATMAAAGAAACGTCGATCCAGTCTTTGATATG 832

RESULT 11
US-09-640-211A-1960
; Sequence 1960, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1960
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1960
```

```
Alignment Scores:
Pred. No.: 1.83e-16 Length: 599
Score: 306.00 Matches: 73
Percent Similarity: 55.11% Conservative: 24
Best Local Similarity: 41.48% Mismatches: 27
Query Match: 18.65% Indels: 52
DB: 4 Gaps: 5

US-10-630-636-7 (1-306) x US-09-640-211A-1960 (1-599)
QY 55 SerValProGlyAlaArg-SerAlaGluGluValArgArgHis-TyrGluAlaLeuValG 74
Db GCGCTGCGCGGTGCGAGATGCTTCCCCACGCGCGCCGACCCCGAGCTGCGCGCGG 189
QY 74 LuAspValAlaAlaAlaAspAlaGly----- 84
Db ACAGAGAGCTCGGCGCGGAGCGCGCGCGGAGATCATGCTTCTGCGGCTGCGGCTG 249
QY 85 -----ProLeuProArgTyrAlaGlyGluG 93
Db TCGTGACCTGATGAGAGAGAGTGGCTGAGCCTGAACAACCTGTCTCAATACAGACCCG 309
QY 93 LuSerAlaAlaProProAspGlyAlaGlyValAlaAlaAlaAlaSerLysAspGlyGly 113
Db AGAGCGGAATCCGCCCCAACGCCAGCGCGCGAGC-----GGCGCA 351
QY 113 isArgArgAspGluArgLysGly-----GlyG 122
Db ACAAGAGAGAGGCGCGCAAGGCTACGCGCCGACGACGCGCGCGCAACCCGCGCG 411
QY 122 LysGlyGlyTyrAspGlyGlyLysSerCysSerLysAlaGluGlnGluArgLysGly 142
Db GTGCG-----CCGAGCGCGAAGAGAG 435
QY 142 LepProTyrThrGluGluGluHisArgLeuPheLeuGlyLeuAspLysPheGlyLys 162
Db TTCCTTGACAGAGAGAGACACAGCTGTTCTTGTTGGATTAACAGAGTGGGAAAG 495
QY 162 LysAspTyrAspSerIleSerArgAsnPheValIleSerArgThrProThrGlnValAla 182
Db GAGATTGAGAGAGGATATCCAGGAATTTGTGAAGACCGCAGCGCCACTCAGGTGCGA 555
QY 182 eHisAlaGlnLysTyrPheIleArgLeuAsnSerMetAsnArg 196
Db 556 GCCATGCCCGAAATATTTCTGCGCGAAGCAACTCAATCCG 599

RESULT 12
US-09-640-211A-1984
; Sequence 1984, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640.211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1984
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1984

Alignment Scores:
Pred. No.: 1.35e-16 Length: 332
Score: 303.50 Matches: 56
Percent Similarity: 81.61% Conservative: 15
```



```
QY 166 SerIleSerArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGln 185
Db 469 GGCATTTCGAGAGACTTTGTAAGACTGCGACCGACCGAGTCGCGAGTCAAGCCCAA 528
QY 186 LysTyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerSerIleHisAsp 205
Db 529 AAATACCTTTCCTCCCGGAGACACCTTAATCG----- 561
QY 206 IleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGly 225
Db 561 ----- 561
QY 226 HisGlnAlaThrGlyAsnProAlaAlaAlaLeuGlyProProGlyMetLysHis 245
Db 562 -----CGGTCCGCGCC-----GGACTTAGCCTTGTT 588
QY 246 HisHisHisPro 250
Db 589 TGACATCACCACCCG 603
```

```
RESULT 15
US-09-640-211A-2059
; Sequence 2059, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2059
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-2059
```

```
Alignment Scores:
Pred. No.: 1.86e-15 Length: 624
Score: 294.00 Matches: 62
Percent Similarity: 67.92% Conservative: 10
Best Local Similarity: 58.49% Mismatches: 20
Query Match: 17.92% Indels: 14
DB: 4 Gaps: 3
```

```
US-10-630-636-7 (1-306) x US-09-640-211A-2059 (1-624)
QY 95 AlaAlaProProAspGlyAlaGlyAlaAlaAlaAlaSerLysAspGlyGlyHisArg 114
Db 336 GCTGACCCCTCCGAGCGCGCTCTGCGCTGCTGCGCG----- 377
QY 115 ArgAspGlnArgLysGlyGlyGlyGlyTyr-----AspGly-----GlyLysSer 130
Db 378 -----GCGGCGCGCGCGCAGTGGCTATCTCTGTATGTCTTGTGAAGCCTCC 425
QY 131 CysSerLysAlaGlnGlnArgArgLysGlyLysProTPrThrGlnGlnGlnHisArg 150
Db 426 ACTTCCCTCCCAATTCCTCGAGCGGAGAAAGGTGTGCTGATGACAGAGAGACATAGA 485
QY 151 LeuPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTrpAspSerIleSerArgAsn 170
Db 486 ATGTTTTCCTAGAGTTTGCAAGAACTTGCAAGAGTGTGAGAGAGATTCACGGAAT 545
QY 171 PheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLysTyrPheIleArg 190
Db 546 TTGTGATTAACAGACAGACCTACACAGGTAGCCAGCCATGACAGAGAATATTATTATCGA 605
QY 191 LeuAsnSerMetAsnArg 196
```

```
Db 606 CAGAGCAATATGACTAGA 623
```

```
Search completed: July 30, 2005, 01:28:20
Job time : 224 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 29, 2005, 22:49:43 ; Search time 716 Seconds

(without alignments)  
2766.119 Million cell updates/sec

Title: US-10-630-636-7

Perfect score: 1641

Sequence: 1 MTSQAAATTTTAAAAAAWTR.....GHAPVVPVGPAPPAKMHQ 306

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgn2\_1/USPFO\_epool/US10630636/runat\_27072005\_123229\_12221/app\_query.fasta\_1.455  
-DB=Published Applications\_NA -OFMT=faetap -SUFFIX=trpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10630636@cgn2\_1.1.789@runat\_27072005\_123229\_12221  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEBUERT -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PC1\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1641	100.0	918	19	US-10-630-636-4
2	1641	100.0	1330	19	US-10-630-636-1
3	1153.5	100.0	1433	19	US-10-437-963-59484
4	1153.5	70.3	1299	20	US-10-425-115-168067
5	1087	66.2	1516	20	US-10-739-930-2650
6	727	44.3	1281	18	US-10-425-114-11058
7	727	44.3	1318	18	US-10-424-599-78776
8	707.5	43.1	1523	19	US-10-437-963-27578
9	698.5	42.6	823	18	US-10-425-114-4609
10	645	39.3	2366	18	US-10-424-599-57473
11	634	38.6	965	19	US-10-767-701-9031
12	628	38.3	1230	18	US-10-425-114-7669
13	624.5	38.1	1432	14	US-10-286-264-35
14	624.5	38.1	1432	15	US-10-286-264-187
15	624.5	38.1	1432	17	US-10-302-267-157
16	624.5	38.1	1432	18	US-10-412-6998-863
17	571	34.8	800	20	US-10-425-115-113945
18	525.5	32.0	3091	20	US-10-425-115-113947
19	516.5	31.5	1896	20	US-10-425-115-184212
20	497.5	30.3	1140	19	US-10-437-963-34443
21	494	30.1	1613	20	US-10-425-115-52609
22	492	30.0	1202	18	US-10-425-114-29755
23	492	30.0	1363	18	US-10-424-599-34323
24	489	29.8	1416	18	US-10-424-599-34322
25	488	29.7	1893	20	US-10-425-115-52607
26	484	29.5	2481	20	US-10-425-115-440
27	482	29.4	1388	20	US-10-425-115-52610
28	473	29.3	1349	19	US-10-437-963-96811
29	473	28.8	1021	18	US-10-425-114-7450
30	473	28.8	1039	18	US-10-424-599-8016
31	470	28.6	1793	20	US-10-425-115-183322
32	469	28.6	1028	20	US-10-425-115-183317
33	457.5	27.9	867	10	US-09-934-455-253
34	457.5	27.9	867	17	US-10-225-066A-653
35	457.5	27.9	867	17	US-10-225-067-119
36	457.5	27.9	867	17	US-10-374-780A-2679
37	457.5	27.9	867	22	US-10-225-066A-653
38	456	27.8	668	10	US-09-934-455-249
39	456	27.8	668	17	US-10-225-066A-339
40	456	27.8	668	17	US-10-374-780A-2677
41	456	27.8	668	22	US-10-225-066A-339
42	455.5	27.8	608	19	US-10-021-323-528
43	443.5	27.0	1523	19	US-10-437-963-56655
44	441	26.9	548	20	US-10-425-115-168064
45	437	26.6	1136	18	US-10-425-114-7392

#### ALIGNMENTS

RESULT 1  
US-10-630-636-4  
; Sequence 4, Application US/10630636  
; Publication No. US20040107456A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu. Su-May  
; TITLE OF INVENTION: PLANT MYB GENES  
; FILE REFERENCE: 08919-088001  
; CURRENT FILING DATE: 2003-07-30  
; PRIOR APPLICATION NUMBER: US 60/399,999  
; PRIOR FILING DATE: 2002-07-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: FastSeq For Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 918







QY 177 TOTHGIVAlAlaSerHISalagInLYTYrPheIIeaRgLeuSenSerMetAsnArgA 197  
DB 811 CGAGCCAGGTGGCGAGCCAGCGGAGAAATATTATTCGGCTCAACTGATGAACCGGG 870  
QY 197 APHATGATGSeSerIleHISaPrlIeThSeValThraGlyAerGlnValAla 217  
DB 871 ACCGGCGCGCTCCAGCATCCAGCATCCAGCGGTGACCCCGCGGAGTGGCGCGG 930  
QY 217 IagGInGInGlyAlaProIleThrgLYHISGInAlaThSeGlyAsnProAlaAlaAla 237  
DB 931 CC-----GGCGCGCCATCCAGCGGT- KCGGCCCGACCGCGGGGGGCAATGCCGA 983  
QY 237 eUGlyProProGlyMetLYsHISaPrlIeThSeValThraGlyAerProProM 257  
DB 984 TGGGCGCGCGCGCGCATGAAGCACACAC-----CCGCTCCGCGCA 1025  
QY 257 eTPromETySeSerAlaAPromETGlyHISProValAlaGlyHISMetVal---ProA 276  
DB 1026 TGGGCGATGACCGGCATGGCCCATGGCCACCGCGCTCGGCGACATGGTGGCGCGG 1085  
QY 276 IAlaValAlGlyThrProValAlaPheProProGly---HISAlaProTyValAlaProV 295  
DB 1086 CGGCGGTGGCGACGCGCGGTATGTCGCGCGGCGCATCATCTGCTTATGTCGCGG 1145  
QY 295 aIGlyTYrProAlaProProAlaLYsMetHISGln 306  
DB 1146 TGGGATACCGGCGCCGACCGACCAAGATGCACAG 1180

RESULT 6  
US-10-425-114-11058  
; Sequence 11058, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 11058  
; LENGTH: 1281  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701002166\_FLI  
US-10-425-114-11058

Alignment Scores:  
Pred. No.: 2,576-61 Length: 1281  
Score: 727.00 Matches: 164  
Percent Similarity: 62.58% Conservative: 40  
Best Local Similarity: 50.31% Mismatches: 70  
Query Match: 44.30% Indels: 52  
DB: 18 Gaps: 13

US-10-630-636-7 (1-306) x US-10-425-114-11058 (1-1281)

QY 10 ThThraAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 29  
DB 72 ACCATGTCATCAAGTGAACCATTTGGAGCTATGAGAGAAAGAAAGCATTTGAGATGCC 131  
QY 30 LeuAlaAlaCySaAlaAlaAlaProProProAlaAlaProGlyAlaProAlaProAlaPro 48  
DB 132 ATAGCTAAG-----CATTGAGATT 149  
QY 49 -----PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSer 61

DB 150 GAGGAATCCTCAAAAGACCAATGGAGAAATTGCTTCAGCGATTCCC---AGCAAAAGC 206  
QY 62 AlAGInGInValAlaArgHISerGlyAlaLeuValAlaSerValAlaAlaAlaSerAla 81  
DB 207 ATGGAAGATGGAACCAATTACGATTTAGTAGAGAGATGATGATGATGATGATGATGATG 266  
QY 82 GATGValProLeuProAlaTYrAlaGlyGlu-----SerAlaAlaProProAla 99  
DB 267 GGTCAATATGATTCCTCAAACTATGCTCTTGAGAAATACATCTTCAAAATAGAGC 326  
QY 100 GlyAlaGlyAlaAlaAlaAlaAlaSerLYsAerGlyGlyHISaRgAerGlyAlaRgLY 119  
DB 327 TTTCAATGCTCTTCAGAGCCCAAGATCAGATAAAGATCAAAATTTGATATGAGAGCT 386  
QY 120 GLYGlyGlyGly---GLTYrAsP-----GLYGLYsSer 130  
DB 387 GGTTCCTGAGGTATGACATGATCCACCACTCATAGTACTGTAAGAGGCTTGTCA 446  
QY 131 CysSerLYsAlaGInGInGlyAerGlyGlyIleProTYrThrgInGlyGlyHISaRg 150  
DB 447 AGCTCA---TCAGAACAAAGAAAGAAAGAAATCCATGATGATGATGATGATGATGATG 503  
QY 151 LeuPheLeuGlyLYsAerLYsPheGlyLYsGlyAerTPArGSerIleSerArgAn 170  
DB 504 TTATTTTACTTGGTCTAGAAAGATTGGAAAGAGATGGAGAAAGCATTTCAAGAAC 563  
QY 171 PheValIleSerAlaThrProThrgInValAlaSerHISaAlaGInLYTYrPheIIeaRg 190  
DB 564 TTTGTGATATCTAGACATCCCATCACTCAAGTGCAGGCCATGCAAAAGTACTTCTAAAG 623  
QY 191 LeuAsnSerMetAsnArgAspArgArgArgSerSerIleHISaPrlIeThSeValThr 210  
DB 624 TTGAATTCAATGATAGAGACAGAGAGGCTTACTATTCATGATATCACTAGTGTGAAC 683  
QY 211 AlAGlyAerGInValAlaAlaGInGInGlyAlaProIleThrgLYHISGInAlaThrgLY 230  
DB 684 AATGAGAT-----GTGGTAAATGAACCAAGCACCTATTACAGGGCAGCATGTAGCACA 737  
QY 231 AsnProAlaAlaAlaAlaLeuGlyProProGlyMetLYsHISaPrlIeThSeValPro 250  
DB 738 ATTCCTTAAACACAAATGGGTAGACAAATCCCTTAAGCATAGATTCAAGGCTCA--- 794  
QY 251 GlyGlyAlaProProPromETySeSerAlaAlaAPromETGlyHISProValAla 270  
DB 795 -----ATACCACTGCTTGGCATGAT---GGAACACAGATTGACATCTGTGGCT 845  
QY 271 -----GLYHISMetValProAlaAlaValGlyThrProValAlaPheProProGly 287  
DB 846 GCTCTCCAGGACATGCA---TCTGCAATGGCACTCCGTATCATCTCTCTGGA 902  
QY 288 -----HISAlaProTyValAlaProValGlyTYrProAlaPro 300  
DB 903 CCCCACCCCATGTCATCTCATCTTCACTTATGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 962  
QY 301 ProAlaLYsMetHISGln 306  
DB 963 CTTCCACAAATGCATCA 980

RESULT 7  
US-10-424-599-78776  
; Sequence 78776, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28

```
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 78776
; LENGTH: 1318
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4214C.1
US-10-424-599-78776

Alignment Scores:
Pred. No.: 2,656-61      Length: 1318
Score: 727.00           Matches: 164
Percent Similarity: 62.58%      Conservative: 40
Best Local Similarity: 50.31%    Mismatches: 70
Query Match: 44.30%           Indels: 52
DB: 18                   Gaps: 13

US-10-630-636-7 (1-306) x US-10-424-599-78776 (1-1318)

Qy 10 ThrThraAlaAlaAlaAlaAlaAlaTThraGluAspAlaPheGluAsnAla 29
Db 72 ACCATGTCATCAAGTGAACCATTTGGAGCTATGAGAGAGAAAAGCATTTGAGAAATGCC 131
Qy 30 LeuAlaAlaCysAlaAlaProProAlaAspGlyAlaProAspAspTyr--- 48
Db 132 ATAGCTATG-----PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSer 61
Qy 49 -----GAGAAATCTCAAAAGAGCATGGAGAGAAATGCTTCAGACAGTTCC--AGCAAAAGC 206
Db 150 GAGAAATCTCAAAAGAGCATGGAGAGAAATGCTTCAGACAGTTCC--AGCAAAAGC 206
Qy 62 AlaGluGluValAlaGluGluHisTyrGluAlaLeuValGluAspValAlaAlaAlaIleAspAla 81
Db 207 ATGGAAGATGGAAGCAACATTAACAGGTTCTAGTAGAGATGTAAGGCAATAGAGCA 266
Qy 82 GlyArgValProLeuProArgTyrAlaGlyGluGlu-----SerAlaAlaProProAsp 99
Db 267 GGTCACTATCATCTCCAAATCTATGCTCTCTGAGAAATCACTCTCAAAATAGAGAC 326
Qy 100 GlyAlaGlyAlaAlaAlaAlaAlaSerLysAspGlyGlyHisArgAspGluArgLys 119
Db 327 TTTCATGGCTCTTCCAAAGGCCACCAAGCTCAGATAAAGATCAAAATTTGTAATTATGACGT 386
Qy 120 GlyGlyGlyGly---GlyTyrAsp-----GlyGlyLysSer 130
Db 387 GGTCTTCTGGGTAGGACATGACTCCACCATATGATGTGTAAAGAGGCTTGCA 446
Qy 131 CysSerLysAlaGluGluGluArgArgGlyLysProTyrPThrGluGluHisArg 150
Db 447 AGGTCA--TCGAACCAAGAAAGAAAGAAAGAAATCCCATGACTGAAGAGAAACACAGG 503
Qy 151 LeuPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTyrArgSerLysSerArgAsn 170
Db 504 TTATTTTACTTGCTGTAGAAAAGTTTGGAAAAGAGATTGGAAGAGATTTCAGAGAC 563
Qy 171 PheValIleSerArgThrProThrGluValAlaSerHisAlaGlnLysTyrPheIleArg 190
Db 564 TTTGTGATATCTAGACACTCCACTCACTGCAAGCCATGCACAAAAGTACTTCATAGAG 623
Qy 191 LeuAsnSerMetAsnArgAspArgArgArgSerSerLysAspLysIleThrSerValThr 210
Db 624 TTGAATTCATGAATAGAGACAGAGAGGCTTATGATTCATGATATCATCTAGTGTGAC 683
Qy 211 AlaGlyAspGluValAlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGly 230
Db 684 AATGAGAGT-----GTGGCTAATTAACCAACACCTATTATACAGGAGCAGCATAGAGCA 737
Qy 231 AsnProAlaAlaAlaAlaLeuGlyProProGlyMetLysHisHisHisHisHisPro 250
Db 738 ATTCCTTCAACACAAATGGGTGTAGACATCCCTTAAGCATTAAGTTCAAGGTCAC-- 794
Qy 251 GlyGlyAlaProProProMetProMetTyrSerAlaAlaProMetGlyHisProValAla 270
Db 795 -----ATACCACCTGCTTAGCAGATAT---GGAACACACAGTTGACATCTGTGGCT 845
Qy 271 -----GlyHisMetValProAlaAlaValGlyThrProValValPheProGly 287
Db 846 GCTCTCTCAAGGACATGAGCA--TTCGAGTTGGCACTCTGTCATGCTCTCTCTGGA 902
Qy 288 -----HisAlaProTyrValValProValGlyTyrProAlaPro 300
Db 903 CCCCACCCCAGTCTCATCTCATTCACCTTAATGTTCTCTCTGTTACCAATGCA 962
Qy 301 ProAlaLysMetHisGln 306
Db 963 CTTCAACATGATCA 980

RESULT 8
US-10-437-963-27578
; Sequence 27578, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 27578
; LENGTH: 1523
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32258C.1
US-10-437-963-27578

Alignment Scores:
Pred. No.: 2,486-59      Length: 1523
Score: 707.50           Matches: 167
Percent Similarity: 58.57%      Conservative: 21
Best Local Similarity: 52.02%    Mismatches: 76
Query Match: 43.11%           Indels: 57
DB: 19                   Gaps: 13

US-10-630-636-7 (1-306) x US-10-437-963-27578 (1-1523)

Qy 12 AlaAlaAlaAlaAlaAlaAlaTThraGluAspAlaPheGluAsnAlaLeuAla 31
Db 466 TCCGCTGGCGGCGGTGTGAGCGCGGAGAGAAAGCGTTGAGAAAGCGCTGGCGG 525
Qy 32 AlaCysAlaAlaProProProAlaAspGlyAlaAlaProAspAspAspTyrPheAlaAla 51
Db 526 ACGGTG-----GGGATGACGAGAGAGAGAGAGGCGGCTTGTGGAGAG 570
Qy 52 LeuAlaAlaSerValProGlyAlaArgSerAlaGluGluValAlaArgArgHisTyrGluAla 71
Db 571 CTACGAGAGCGCGTGGAGGG--AGACGCGCAGACAGAGGTGAGCGGCACTACGACTG 627
Qy 72 LeuValGluAspValAlaAlaAlaIleAspAlaGlyArgValProLeuProArgTyrAlaGly 91
Db 628 CTGGTGAGAGCGTGCAGCGGATCGAGCGCGGCGGCGGTCCCTGCTGTAGCGCGG 687
Qy 92 GluGluSerAlaAlaProProAspGlyAla-----GlyAlaAlaAlaAlaAlaSer 108
Db 688 -----GACGGGGCGCTCGAGAGAGGCGCTTGTGGGAGGTGGAGAG 726
Qy 109 LysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGlyGlyTyrAspGlyGly 128
Db 128
```

```

Db      727 AAGGGGGTGT-----GGGGGAGAGGTGGAGCTGAGGGGG 765
Qy      129 Lys-----SerCysSerIysAlaGlnGlnAlaArgLysGlyIleProTyr 144
Db      766 CATGGGAGAAAGGGTCCGCTAAGTCCGCTGAGCAGAGCGCCGGAAGGGGATCCGCTGG 825
Qy      145 ThrGlnGlnGlnHisArgLeuPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTyr 164
Db      826 ACGGAGACAGACAGACAGGCTGTTCTCTTGGACTTGAAGATACGGCAAGAGCCAGCTGG 885
Qy      165 ArgSerIleSerArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAla 184
Db      886 AGGACTATCTCAAGAACTTGTGATCTCAGAGAACCCCACTGAGCTGTGATGCA 945
Qy      185 GlnLysTyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSerSerIleHis 204
Db      946 CAGAGATATTTTATTCGCTGATCATGACAGAGAGGCGCGCATCAAGTACAT 1005
Qy      205 AspIleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThr 224
Db      1006 GACATACACAGCGTGAACATGAGATACATGCTGCTGCTCAG-----GGCCCAATCACA 1059
Qy      225 GlyHisGlnAlaThrGlyAsnProAlaAlaAlaAlaLeuGlyProProGlyIleLysHis 244
Db      1060 GGT---CAGCCAAATGGC---CCATCAGCA-----AATCTGGAAATCTCT 1101
Qy      245 HisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis 1158
Db      1102 AAGCAGTCTCTACAGCCAGCAATGCGCTCAGGCGTGTGATGCTTATGTCAGCA--- 1158
Qy      265 MetGlyHisProValAlaGlyHisMetValProAlaAlaValAlaGlyThrProValAlaPhe 284
Db      1159 ATTGACAGCAGCGTGTGTGCTCTCTTG---TCGAGGTGGACCTCTTACACT 1215
Qy      285 ProProGlyHisAlaProTyrValVal----- 293
Db      1216 CCGTCTCTGTGCTGACCTCATATGACCTATGCGATGATGCGCCCTGAGACTGTA 1275
Qy      294 -----ProValGlyTyrProAlaProProAlaLysMet 304
Db      1276 GTCCCTGTGCCCCAGTAAACATGCTCCATGCTTACCCCATCGCCACCAACATCT 1335
Qy      305 His 305
Db      1336 CAT 1338

RESULT 9
US-10-425-114-4609
; Sequence 4609, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-26
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4609
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700397574_FLI
US-10-425-114-4609
Alignment Scores:
Pred. No.: 9,686-59 Length: 823

```

```

Score: 698.50 Matches: 139
Percent Similarity: 82.86% Conservatve: 6
Best Local Similarity: 79.43% Mismatches: 13
Query Match: 42.57% Indels: 17
DB: 18 Gaps: 6

US-10-630-636-7 (1-306) x US-10-425-114-4609 (1-823)
Qy      137 GluArgArgLysGlyIleProThrGlnGlnGlnHisArgLeuPheLeuLeuGlyLeu 156
Db      1 GAGCGCCGAAAGGGGCTCCGTGAGCAGAGAGACAGAGTGTCTTGTGGGCTTG 60
Qy      157 AspLysPheGlyLysGlyAspTyrPargSerIleSerArgAsnPheValIleSerArgThr 176
Db      61 GACAACTTCGGAAGAGGGGACCTGGGAGCATCTCGGCCAATTGTCATCTCGGGAGC 120
Qy      177 ProThrGlnValAlaSerHisAlaGlnLysTyrPheIleArgLeuAsnSerMetAsnArg 196
Db      121 CCGACGCGAGTGGCCAGCCAGCGCGCAAAATCTTCATCCGCTCACTCGATGAAACCGG 180
Qy      197 AspArgArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGlnValAla 216
Db      181 GACAGCGCCGCTCCAGATCCACGATCCACAGCGTGAACCGCGCGAG-----GTG 234
Qy      217 AlaGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGlyAsnProAlaAlaAla 236
Db      235 GCCACGGCGCGCGCCCATCACGGC-----GGCGGCGCAGACGGCGGGGG 282
Qy      237 -----LeuGlyProProGlyIleLysHisHisHisHisHisHisHisHisHisHisHis 253
Db      283 GCGATGCCAGTGGGCGCGCGCGCGCATGAAAGCACCCAC-----CCG 324
Qy      254 ProProMetProMetTyrSerAlaAlaProMetGlyHisProValAlaGlyHisMet 273
Db      325 GTTCCGCGATGGGCGATGACGCGCACCGCCCATAGGGCCACCGGTGCGGGGACATG 384
Qy      274 Val--ProAlaAlaValAlaGlyThrProValAlaPheProGly---HisAlaProTyr 291
Db      385 GTGGCGCGGGGGCGGTGGCACCGCGGTCAATGTTCCGGCGGGGCATCATCGCCCTAT 444
Qy      292 ValValProValGlyTyrProAlaProProAlaLysMetHisGln 306
Db      445 GTCTGCGCGTGGGATACCGCGCGCCACCGCAAGATGACACAG 489

RESULT 10
US-10-424-599-57473
; Sequence 57473, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 57473
; LENGTH: 2366
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22910C.1
US-10-424-599-57473
Alignment Scores:
Pred. No.: 5,076-53 Length: 2366
Score: 645.00 Matches: 153
Percent Similarity: 58.49% Conservatve: 33
Best Local Similarity: 48.11% Mismatches: 58
Query Match: 39.31% Indels: 74

```

DB: 18 Gaps: 15

US-10-630-636-7 (1-306) x US-10-424-599-57473 (1-2366)

QY 18 TrpThrArgGluAspAlaPheGluAsnAlaLeuAlaIaIaCysAlaAlaProPro 37

DB 836 TGGGAAAAAGAACGAGATTAAGCATTTGGAAAATGCTTGGCAATT----- 880

QY 38 ProAlaAspGlyAlaProAspAsp-----AspTrpPheAlaAlaLeuAlaIaSer 55

DB 881 -----CACCTGAGGATGCTTCAGACGGGTGGAGAGATTGACGCTGAT 925

QY 56 ValProGlyAlaArgSerAlaGluValArgArgHisTyrGlyAlaLeuValGluAsp 75

DB 926 GTACCAAGG--AAACCTTGGAAGAGATTAACCAACCATATGAGCTCTGTGTGAAGAT 982

QY 76 ValAlaAlaIleAspAlaGlyArgValProLeuProArgTyr-----AlaGly 91

DB 983 GTTAGCCAGATCGATCCGGTTATGTGCTTTACCATCTTATATTTCTTCCAGAGGC 1042

QY 92 GluGluSerAlaAlaProProAspGlyAlaGlyAlaAlaAlaAlaSerLysAspGly 111

DB 1043 TCACAAAGCCATGCTAGTGAAGAAGAGAGCTGGC-----AAGAAGGA 1084

QY 112 GlyHisArgArgAspGluArgGlyGlyGlyGlyGlyTyrAspGlyGlyLysSerCys 131

DB 1085 GGCACCTCTGGAAATAGTAATACCAATCT-----AATCATGGAATCAAGAGCT 1132

QY 132 SerLysAlaGluGluArgArgGlyLysLeuProTrpThrGluGluGluHisArgLeu 151

DB 1133 TCACAGATCAAGTCAAGAACGACGAAAGGGTATTGATGACAGACAGATGAACAGGTTA 1192

QY 152 PheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTrpArgSerLysSerArgAspPhe 171

DB 1193 TTCTCCTTGCTGGCTTGAAGATATGGGAAGGTGACGCGGAGATATATCAAGAGACTTT 1252

QY 172 ValIleSerArgTrpProThrGlnValAlaSerHisAlaGlnLysTyrPheIleArgLeu 191

DB 1253 GTGTGTCAAGAACGCGCTACCAAGATGACCAAGCCATGCCCCAAAATCTTCATTCGCTCG 1312

QY 192 AsnSerMetAsnAspArgArgArgSerSerIleHisAspIleThrSerValThrAla 211

DB 1313 AACTCGATGATTAAGACAGAGGCGATCCAGCATACAGATATACCAAGTGTCAACAT 1372

QY 212 GlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAlaThrGly 230

DB 1373 GGAAGAT---GTTTCAGCAGCTCAAGGA---CCAATTACTGCT---CAACAAACGGTTCT 1423

QY 231 -----AsnProAlaAlaAlaLeuGlyProProGlyMetLysHisHisHisHis 248

DB 1424 GCGGATTAATCTGCTGGAAAATCAACCA----- 1456

QY 249 HisProGlyGlyAlaProProProMetPro-----MetTyrSer 261

DB 1457 -----GCCCAACAGCCCCCACTGCTGCTGGGTGAGGATATATGCT 1501

QY 262 AlaAlaProMetGlyHisProValAlaGlyHisMetValProAlaAlaValAlaGlyThrPro 281

DB 1502 GGTCTCACTACATTGACCAACTATAGAGGAGACCCCTCGTA---TCTGAGTTGGACCCCA 1558

QY 282 ValValPhe-----ProProGlyHisAlaProTyr----- 291

DB 1559 GTGATGAACCTTCTCCCGCCAGCACACATGACATATAGTCTCGGAGCACCGGTTCTGGG 1618

QY 292 -----ValValProValGlyTyrProAlaPro 300

DB 1619 GCAAGTTGTTCTGGAGCAACATGAATTAAGTCTTGCCATACCCATAGCCA 1672

RESULT 11

; Sequence 9031, Application US/10767701

; Publication No. US20040172684A1

; GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yina

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5353)B

CURRENT APPLICATION NUMBER: US/10767, 701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 9031

LENGTH: 965

TYPE: DNA

ORGANISM: Sorghum bicolor

OTHER INFORMATION: CLONE ID: SORBI-28MAY03-CLUS73247\_1

US-10-767-701-9031

Alignment Scores:

Pred. No.: 2,29e-52 Length: 965

Score: 634.00 Matches: 143

Percent Similarity: 60.61% Conservative: 17

Best Local Similarity: 54.17% Mismatches: 62

Query Match: 38.63% Indels: 42

DB: 19 Gaps: 10

US-10-630-636-7 (1-306) x US-10-767-701-9031 (1-965)

QY 63 GluGluValAlaArgArgHisTyrGluAlaLeuValGluAspValAlaAlaIleAspAlaGly 82

DB 960 GAGAGGTGACGCGCGCATACGAGCTGCTGTGAGAGACCTTCACGCGCATCGAGTGGCG 901

QY 83 ArgValProLeuProArgTyrAlaGlyGluGluSerAlaAlaProProAspGlyAlaGly 102

DB 900 CGCGTCCGCTTCCGACGTACGCGGCTGACGCGGCCCGCCGAGAGGGGGTGGCGCGCG 841

QY 103 AlaAlaAlaAlaAlaSerLysAspGlyGlyHisArgArgAspGluArgLysGlyGly 122

DB 840 GGC-----GCGGG-----AAGAGGAGAGTGGGGGA 814

QY 123 GlyGlyTyrAspGlyGlyLysSerCysSerLys---AlaGluGluGluArgArgLysGly 141

DB 813 GAGAGGACCCAGCGGAGCAAGGGGTGCGGAGTGGCGGACGAGCAGCGCAGGAGAGGGC 754

QY 142 IleProTrpThrGluGluGluHisArgLeuPheLeuLeuGlyLeuAspLysPheGlyLys 161

DB 753 ATCCCTCGACAGAGGACGAGCAGAGATTTCTCTCTGAGACTTGAAAAGTATGGCAA 694

QY 162 GlyAspTrpArgSerIleSerArgAsnPheValIleSerArgTrpProThrGlnValAla 181

DB 693 GTGACTGGGAGACATCTACGTAATTTGTGATCTCAAGACTCCAACTCAAGTACGCT 634

QY 182 SerHisAlaGlnLysTyrPheIleArgLeuAsnSerMetAsnArgAspArgArgSer 201

DB 633 AGTATGACAGAGAATCTTATCCGTGATCAATGAACAGAGAGGCGCGCATCA 574

QY 202 SerIleHisAspIleThrSerValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAla 221

DB 573 AGTATCATGATCATCAACAGTGTGAACAAATGAGATCATCAACGCGCAG-----GCT 520

QY 222 ProIleThrGlyHisGlnAlaThrGlyAsnProAlaAlaAlaAlaLeuGlyProProGly 241

DB 519 CCAATCAAGT---CAACAAATGTCAAGCTGCA-----AATCTGGG 478

QY 242 MetLysHisHisHisHisHisHisProGlyAlaAlaProProMetProMetTyrSer 261

DB 477 AAACCATTAAGCAATCCCAACAGCCAGCAATATCACTCCAGCGCTGATAGGC 418

QY 262 AlaAlaProMetGlyHisProValAlaGlyHisMetValProAlaAlaValAlaGlyThrPro 281

DB 417 ACAACA---ATTGACAGCGAGTTGGTGGCCCTTGTCT---TCAGCTGTGGAATCTCT 364

QY 282 ValValPhePro-----ProGlyHisAlaProTyr----- 291









QY 257 MetProMetTyrSerAlaAlaProMetGlyHisProValAlaGlyHisMetValProAla 276  
Db 967 CCTCTATGTATGGAACACCCGCCATAGTCAAGCA----- 1002  
QY 277 AlaValGlyThrProValValPhePro---ProGlyHisAlaProTyrValVal----- 293  
Db 1003 GCAGTTGGAACACCACTGAACTCCCACTCCACCTCACTGCTTATGGAATTCAATGCG 1062  
QY 294 ---ProVal---GlyTyrProAlaProProAlaLysMetHis 305  
Db 1063 GCTCCAGTCCCTGCTCACTGCTTCTGTCAGACAAATGAAC 1104

Search completed: July 30, 2005, 01:40:17  
Job time : 729 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 29, 2005, 22:29:33 ; Search time 3685 Seconds  
(without alignments)  
3160.832 Million cell updates/sec

Title: US-10-630-636-7  
Perfect score: 1641  
Sequence: 1 MTSQAATTTTAAATAAATWTR.....GHAPVYVGVGPAPAKKHQ 306

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgr2\_1/USPRO\_epool/US10630666/runat\_27072005\_123227\_12178/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=plco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10630666 @CGN 1.1 4352 @runat\_27072005\_123227\_12178 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOBJECT -NEG\_SCORES=0 -WAIT -DSPELCK=100 -LONGIDCG  
-DEV\_TIMEOUT=10 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gse1: \*  
9: gb\_gse2: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828.5	50.5	884	9	CG257740 OGXCQ48TH
2	816.5	49.8	689	5	BQ483326 WHB3514.C
3	804	49.0	478	7	CF326660 JMT1--06-
4	787.5	48.0	759	8	BZ525067 OGART807C
5	777.5	47.4	815	6	CA185031 SCSST309
6	762	46.4	853	9	CG362076 OGICL41TH
7	742	45.2	970	7	CK256709 EST740346
8	742	45.2	973	7	CK254513 EST738150
9	737	44.9	1020	7	CK251993 EST735630

10	717	43.7	952	7	CK294026
11	713.5	43.5	846	7	CO082313
12	694.5	42.3	921	7	CK248989 EST732626
13	664.5	40.5	936	7	CK292080 EST754794
14	655.5	39.9	956	7	CK259271
15	655.5	39.9	961	7	CK257717
16	653	39.8	921	9	CG257748
17	635.5	38.7	874	9	CG219919
18	633.5	38.6	1641	3	CNS0A165
19	633.5	38.5	864	7	CO089242
20	627.5	38.2	1554	3	CNS0A05A
21	624	38.0	769	7	CK258863
22	623.5	38.0	689	9	CG710311
23	623.5	38.0	781	7	CV500657
24	623.5	38.0	804	9	CG366978
25	623.5	38.0	933	9	CG219911
26	611.5	37.3	740	9	CG751932
27	609.5	37.1	1672	3	CNS0A13X
28	607	37.0	974	7	CK290192
29	606.5	37.0	653	9	CG752461
30	606	36.9	628	2	BR356185
31	606	36.9	705	6	CA243768
32	603	36.7	831	7	CO121658
33	597	36.4	621	6	CB003696
34	596	36.3	800	7	CK257718
35	593.5	36.2	695	7	CK259272
36	590	36.0	711	5	BU025086
37	584.5	35.6	599	2	BR433827
38	582.5	35.5	643	6	CA233892
39	578	35.2	778	7	CO081265
40	576	35.1	815	7	CF450963
41	572	34.9	841	7	CO119900
42	571.5	34.8	633	5	BO864479
43	569.5	34.7	887	7	CO117508
44	560	34.1	678	6	CB001319
45	560	34.1	678	6	CB002201

#### ALIGNMENTS

RESULT 1  
LOCUS CG257740/C  
DEFINITION OGXCQ48TH ZM 0.7 1.5 KB zea mays genomic clone ZMWBM065G23,  
genomic survey sequence.  
ACCESSION CG257740  
VERSION CG257740.1 GI:34164524  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE  
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
1 (bases 1 to 884)  
Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.  
TITLE JOURNAL  
COMMENT Unpublished (2002)  
Other GSSs: OGXCQ48TV  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source  
1..884  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"



Db		180	TGACAGGAGAGGTCCGGCGAACATCAGCGCCCTGTGAGAGA	GSTCGGCAGTTCAC	239
Oy		81	AAlaGlyAArVaaProLeuProArgTrpAlaGlyUGlUGlUSeRAlAAlAProProAspGly		100
Db		240	GCCCCGCCGCTCCCTCCCGCGCATCCGCCGGGAGGAATCTTCGGCC-----		287
Oy		101	AlaGlyAlaAlaAlaAlaAlaSerIysAspGlyGlyHisSarArgASpGluArgrgylgs		120
Db		288	GCGGGCGCGCGCGCGCGCTTGCAAGGAC-----CGCGCGAGGACCGGAAAGAC		338
Oy		121	GlyGlyGlyGlyTyrrAspGlyGlyIlysSerCyserIysAlaGluGlnGluArgrgyls		140
Db		339	TTGCACTCC-----GACTCCGGAGAGGCGCTCTCCAAGCGCGAGAGAGCGGCGCAAG		392
Oy		141	GlyIleProTrpThrGluGluGluHisSarGleuPheLeuLeuGlyLeuAspIysPheGly		160
Db		393	GGCATCCCTCGACGTAGAGAGAGACAAGCGCTGTTCTTGCGGGCTGGACAAATTCCGCG		452
Oy		161	LysGIyaPTPrARgSerIleserArganPheValIleserArgThProthrGlnVal		180
Db		453	AAGGCGAGCTGGCGGAGCATCTCGCGCANCTTCATCTCCCGAGCGCCACGAGGTG		512
Oy		181	AlaSerHiSaIaGlnLysTyrrPheIleArgLeuAnSerMetAsnArgAsparArgArg		200
Db		513	GCGACGACAGCGGCAGAGACTTCATCCGCTCAATCCTCATTAACCGGGACCGCGCGCG		572
Oy		201	SerSerIleHisAprIleThrservalThrAlaGlyAspGlnValAlaAlaGlnGlnGly		220
Db		573	TCCACACATCCACGACATCCACGACATCAACAAC-----GCCGCCACGACCGGTG		620
Oy		221	AlAProlleThGlyHISglInlaThrGlyAanProAlaAlaAlaLeu---GlyPro		239
Db		621	CCGCCCATCACCGGCGCAG-----GCCCGCGCGCGCTGGGGGGCGCG		662
Oy		240	ProGly-----MetLysHisHisHisHisHisHisProGly		251
Db		663	CCCCGGCGCATCATGAAG-----CATCCCGGG		669
RESULT 3					
LOCUS		CFJ26660	478 bp	mRNA	linear EST 18-AUG-2003
DEFINITION		JMT1--06-K20_g1 AJMWT-overexpressing transgenic rice lambda phage cDNA library (JMT1)		Oryza sativa (japonica cultivar-group) clone JMT1--06-K20, mRNA sequence.	
ACCESSION		CFJ26660			
VERSION		CFJ26660.1		GI:33801577	
KEYWORDS		EST.			
SOURCE		Oryza sativa (japonica cultivar-group)			
ORGANISM		Oryza sativa (japonica cultivar-group)			
REFERENCE		Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzeae; Oryza.			
AUTHORS		Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,Y.K., Kim,Y.-K. and Nahm,B.H.			
TITLE		Large-scale Sequencing Analysis of Rice ESTs Unpublished (2003)			
JOURNAL COMMENT		Contact: Nahm B.H. Genomics and Genetics Institute, Greengene Biotech Inc.; Division of Bioscience and Bioinformatics, Myoungji University Yongin, Kyoeongi, Korea Tel: 82 31 330 6193 Fax: 82 31 321 6355 Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.			
FEATURES		location/Qualifiers			
source		1..478			
		/organism="Oryza sativa (japonica cultivar-group)"			
		/mol_type="mRNA"			
		/cultivar="Nackdong"			
		/db_xref="taxon:39947"			
		/clone="JMT1--06-K20"			
		/tissue_type="leaf"			
		/dev_stage="14 days after germination"			

Alignment Scores:	Pred. No.:	8,93e-52	Length:	478
Score:	804.00	Matches:	152	
Percent Similarity:	97.44%	Conservative:	0	
Best Local Similarity:	97.44%	Mismatches:	2	
Query Match:	48.99%	Indels:	2	
DB:	7	Gaps:	0	
US-10-630-636-7 (1-306) x CFJ326660 (1-478)				
QY	153	LeuLeuGlyLeuAspLysPheGlyLysGlyAspTrpArgSerIleSerArgAsnPheVal	172	
DB	10	TTGGTCGGGGCTGGACAAGTTCTGGCAGAGGGGACATCGCGAGCATTCGCGCACTTCGTC	69	
QY	173	IleSerArgThrProThrGlnValAlaSerHisAlaGlnLysTrpPheIleArgLeuAsn	192	
DB	70	ATCTGGCGGAGCCCAAGCAGGTGCGAGCCAGCCGCAAGATCTTATCGGCTCAAC	129	
QY	193	SerMetAsnAlaGAPArgArgArgSerSerIleHisAspIleThrSerValThrAlaGly	212	
DB	130	TCCATGAACCCGACCGCGCGCTCCAGCATCCACGACATCACAGGTGACCGCGGC	189	
QY	213	AspGlnValAlaAlaGlnGlnGlyValAspIleThrGlyHisGlnAlaThrGlyAsnPro	232	
DB	190	GATCAGGTGCGCGGAGCAGAGGGGCCCGCATCCGGCCACGAGCCAGCCGAGCAACCC	249	
QY	233	Ala-AlaAla-AlaLeuGlyProProGlyMetLysHisHisHisHisHisProGlyG	252	
DB	250	GCGGCGCGCGGCGCTGGCGCGCGCGCGGCAAGAACACACACACACACACCGCGGCG	309	
QY	252	LysAlaProProProMetProMetTrpSerAlaAlaProMetGlyHisProValAlaGlyH	272	
DB	310	GCGGCGCGCGCGCATGCGCATGACAGCGCGCGCGCATGAGGCAACCGCTCGCGGCG	369	
QY	272	IleMetValProAlaAlaValGlyThrProValAlaPheProProGlyHisAlaProTrpY	292	
DB	370	ACATGTGTCGGCGCGCGCGTGGCGACGCGCGGTGTTCCGCGCGGCGCGCGCGTACG	429	
QY	292	AlValProValGlyTrpProAlaProProAlaLysMetHisGln	306	
DB	430	TCTGTCGCGCTCGCTACCGCGCGCGCACCGCGCAAGATGCACCAA	473	
RESULT 4				
BZ525067/c				
LOCUS	BZ525067	759 bp	DNA	linear
DEFINITION	OGAHT80TC ZM2_0.7.1.5_KB Zea mays genomic clone ZMMBma0064M16,			
ACCESSION	BZ525067			
VERSION	BZ525067.1			
KEYWORDS	GSS.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
	clade; Panicoidae; Andropogoneae; Zea.			
	1 (bases 1 to 759)			
	WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,			
	Reenick,A., Frazer,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,			
	Citek,R.W., Numbers,A., Robbins,D. and Lakey,N.			
	Consortium for Maize Genomics			
	Unpublished (2002)			
	TIGR			













```

QY 123 -----GlyGlyTyr-----AspGlyGly 128
Db 366 TCTTTTGTAGTATGTAATGATTTTCAGAGTACTGATCAATGAATTCGCCAAGAGGA 425
QY 129 LysSerCySeSerLysAlaGluGlnGluArgArgGlyGlyLeuProTrpThrGluGluGlu 148
Db 426 AAGAGAGTTCGAATTCGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 485
QY 149 HisAlaGluPheLeuLeuGlyLeuAspLysPheGlyLysGlyAspTrpArgSerIleSer 168
Db 486 CATAGAGTTTCTCTGCTGTTAGATGAATGCGAAGGAGGAGCTGGAGGAAGTATTTCA 545
QY 169 ArgAsnPheValIleSerArgThrProThrGlnValAlaSerHisAlaGlnLysTrpPhe 188
Db 546 AGGAATATGATGATTTCTAGAACACCACTCAAGGCTAGTCAAGCCCAAGAAAGTACTT 605
QY 189 IleArgLeuAsnSerMetAsnArgAspArgArgArgSerSerIleHisAspIleThrSer 208
Db 606 ATCCGATTAAATTCATGATAGATAGAGTAGAGAGTCCAGCATGACATTCATCAAGT 665
QY 209 ValThrAlaGlyAspGlnValAlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAla 228
Db 666 GTCAACAAATGGG---GAGGTGCTCAAAATCAG---ACTCTTAATAACAGGCAAAACACC 719
QY 229 ThrGlyAsnProAlaAlaAlaLeuGlyProProGlyMetLysHisHisHisHis 248
Db 720 AATACAGCCCATCAAAATGACCTGAAACCCAGTGTGGGCAAAAGAACAGCCCAATATG 779
QY 249 HisProGlyGlyAlaProProMetProMetLysSerAlaAlaProMetGlyHisPro 268
Db 780 CAT-----GGTTGAGTATGAT---GGTGTCCCATGAGTGTATCCG 818
QY 269 ValAla-----GlyHisMetValProAlaAlaValGlyThrProValAlaPhePro 285
Db 819 GTCCGTGCTCCACCCGGGACCAACCGTG---TCAGCAGTCGGAACACCGGTCTCTCT 875
QY 286 ProGly---HisAlaProTrpValValProValGlyTrpProAlaProPro 301
Db 876 CTGAGCATCATCCACCTTATGTTCTTCCAGCTGATACACATGCGCGG 926

RESULT 11      846 bp      mRNA      linear      EST 15-JUN-2004
LOCUS      CO082313
DEFINITION GR_Ea46N19.r GR_Ea Gossypium raimondii cDNA clone GR_Ea46N19.3',
            mRNA sequence.
ACCESSION      CO082313
VERSION      CO082313.1 GI:48751794
KEYWORDS
SOURCE
ORGANISM      Gossypium raimondii
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
            1 (bases 1 to 846)
REFERENCE
AUTHORS      Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
            Udell, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
            Wing, R.A.
            Global assembly of cotton ESTs
            Unpublished (2004)
            Contact: Rod A. Wing
            Arizona Genomics Institute
            The University of Arizona
            Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            Plate: 46 row: N column: 19.
FEATURES
            source
            1..846
            /organism="Gossypium raimondii"
            /mol_type="mRNA"
            /db_xref="taxon:29730"

```

```

/clone="GR_Ea46N19"
/cisue_type="whole seedlings"
/dev stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"
/notes="vector: pCMVSPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:
Pred. No.:      1,44e-44      Length:      846
Score:          713.50      Matches:      163
Percent Similarity: 63.82%      Conservative: 31
Best Local Similarity: 53.62%      Mismatches: 61
Query Match:    43.48%      Indels:      49
DB:              7      Gaps:      14

US-10-630-636-7 (1-306) x CO082313 (1-846)

QY 13 AlAlaAlaAlaAlaTTPThArgGluAspLysAlaPheGluAsnAlaLeuAla--- 31
Db 2 GCAGAGAGTGTAGTTGAGCAGGAGCAATGAGAGGCAATTGAAACCAATACCAATG 61
QY 32 -----AlaCysAlaAlaProProProAlaAspGlyValaProAsp 45
Db 62 CATTGAGATGAGAAAGAGAAATG-----GAA 88
QY 46 AspAspTrpPheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluGluVal 65
Db 89 GAGCAATGG---GAAAGATTTGCTTCAACGGTCTT---ACTAAAGCTTGGAGAAATGG 142
QY 66 ArgArgHisTyrGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArgValPro 85
Db 143 AAGCTTCACCTAGCAATTTTATGTCGAGGATGTTACTGCATTAAGAGCTGTCATGTTCA 202
QY 86 LeuProAspTyrAlaGlyGluGlu-----SerAlaAlaProProAspGlyValaGlyAla 103
Db 203 CTTCCTGTTTAAAGGAGAAACCTTCTTCTGTAAGATTAATTTTCATGAGACT 262
QY 104 AlAlaAlaAlaSerLysAspGlyGlyHisArgArgAspGluArgGlyGlyGly 123
Db 263 TCCATGGCTCTCAATTCGAT-----AGGAATCCCAATTCGGTTACGGAAT 310
QY 124 GlyTyrAsp-----GlyGlyLysSerCysSerLysAla 134
Db 311 AGGTTTGGGGGTTAACCTTGATTTCTACTGACATGAGGAGAAACAGATTCGAGTCC 370
QY 135 GluGlnGluArgArgGlyGlyLysProTrpThrGluGluGluHisArgLeuPheLeuLeu 154
Db 371 GACCAAGAAAGAAAGAAAGAAATCCCATGACTGAAGAAAGCAATAGGATTTTCTACTT 430
QY 155 GlyLeuAspLysPheGlyLysGlyAspTrpArgSerIleSerArgAsnPheValIleSer 174
Db 431 GCTTTAGCAAGTTTGGAGAAAGAGATTCGAGAACATTTCAAGAACCTTTGTGATACG 490
QY 175 ArgThrProThrGlnValAlaSerHisAlaGlnLysTyrPheIleArgLeuAsnSerMet 194
Db 491 AGAATCCGACGACGAGTGGCTAGCCATCACAAAAGATTTTAATTCGCTTAATTCGATG 550
QY 195 AsnArgAspArgArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGln 214
Db 551 AATAGAGACCGCGCGGTCCAGCATTCACACATCAAGAGGTTTAACAAATGGGAT--- 607
QY 215 ValAlaAlaGlnGlnGlyAlaProIleThrGlyHisGlnAla---ThrGlyAsnProAla 233
Db 608 -----ACGCTCATCAAGCCCTATTCACCGCCCAACAGGCTTAACAACTCACCGGCT 661
QY 234 AlAlaAlaLeuGlyProProGlyMetLysHisHisHisHisHisHisHisProGlyGlyAla 253
Db 662 GCAGCGGTCAATGGG---CAATCATGTAAAGACAGGCGCTACGCGCAT----- 706

```





Percent Similarity: 63.44% Conservative: 29  
 Best Local Similarity: 53.05% Mismatches: 53  
 Query Match: 39.95% Indels: 49  
 Gaps: 12

US-10-630-636-7 (1-306) x CK25717 (1-956)

```

QY 13 ALAAlaAlaAlaATPThrArgGluAspAspLysAlaPheGluAsnAlaLeuAlaAla 32
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 200 TCATCTTCAGTACGTGAGCAGAGAAAGAAAGAAAGAAAGCTTTGAGATGCGATGCTGTG 259
QY 33 CysAlaAlaProProProAlaAspGlyGlyAlaProAspAspAspTrp----- 48
Db 260 -----PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluVal 277
   49 -----PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluVal 65
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 278 GAACAGCAATGAGGAGCACTTGTCTTATGTCCTCC--ACAAAGACCATGATGACCTT 334
QY 66 ArgArgHisTyrGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArgValPro 85
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 335 AAGGAAACATTACCAAGTACTTTTGAAAGATGTCACCACTTGAAGCAGGCCAAGTCCA 394
QY 86 LeuProArgTyrAlaGlyGluGlu---SerAlaAlaProProAspGlyAlaGlyAlaAla 104
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 395 ATACCCCAATTACAAAGAGAAAGAGCTTCTCTTCACTAAAGAAAGTTAATCTAGGA 454
QY 105 ALAAlaAlaSerLysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGly 124
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 455 TATCTGGGTCAAGTACGACGGA-----CGGCGCTCCAAATTGGTTATTCAAAGGA 505
QY 125 TyrAsp-----GlyGlyLysSerCysSerLysAlaGlu 135
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 506 TTTTCAGGTACGACCCATGACCCGATTGGCGATGGGAAAGGAAATTCAGGCTCGAA 565
QY 136 GluGluArgArgLysGlyLysProTyrThrGluGluGluHisArgLysLeuLeuGly 155
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 566 CAAAGAGAGCAAGAAAGCATACCGTGGACTGAAGAAAGCAATAGCTGTTTCTAGGT 625
QY 156 LeuAspLysPheGlyLysGlyAspTyrAspSerLysSerArgAsnPheValIleSerArg 175
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 626 TTACACAAATTTGGGAAAGAGATTTGAGAGATTTCAAGAGATTTGTTATATCTCGA 685
QY 176 ThrProThrGluValAlaSerHisAlaGluLysTyrPheIleArgLeuAsnSerMetAsn 195
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 686 ACACCAACACAAAGTACGTATCATGCCCAAGAAATCTTATTCGTTGAATTCATGAAT 745
QY 196 ArgAspArgArgArgSerSerLysAspLysThrSerValThrAlaGlyAspGluVal 215
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 746 AGAGATGAGAGAGGTCAAGTATTCAGCATTTACAGATTCAAATGAGGAGACGTT 805
QY 216 ALAAlaGluGluGlyAlaProIleThrGlyHisGluAlaThrGlyAsnProAla---Ala 234
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 806 TCACACTCATCA---GCTCCCAATTACAGGCCAACAGAGATTCGAATTCATCAATCCA 862
QY 235 ALAAlaLeuGlyProProGlyMetLysHisHis-----HisHisHisHisProGly 251
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 863 GCTCTCTTTGGA---CCCTCAGTACAGACGAAACCCCAACCAATATGACATAGATGAGGC 919
QY 252 GlyAlaProProProMetProMetTyrSerAlaAlaProMetGlyHisProValAla 270
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 920 -----ATGTAC---GCTGCTCGATGGGTCAACCCCGTGTCT 952

```

RESULT 15  
 CK25717 961 bp mRNA linear EST 30-JUN-2004  
 LOCUS E57741354 potato callus cDNA library, normalized and full-length  
 DEFINITION Solanum tuberosum cDNA clone POC0756 5' end, mRNA sequence.  
 ACCESSION CK25717  
 VERSION CK25717.1 GI:39814697  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum

REFERENCE  
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karanymchava, S.A. and Baker, B.  
 TITLE Generation of ESTs from potato callus tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: E57741355  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr. Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics  
 Institute via <http://genome.arizona.edu/orders/>.  
 Seq primer: ATT TAG GTG ACA CTA TAG.  
 Location/Qualifiers

#### FEATURES

source  
 1. 961  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POCD756"  
 /rissue\_type="callus"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato callus cDNA library, normalized and full-length"  
 /note="Vector: pCMVSPORT6.1; Site\_1: EcoRI; Site\_2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

#### ORIGIN

##### Alignment Scores:

Pred. No.: 4.51e-40 Length: 961  
 Score: 655.50 Matches: 148  
 Percent Similarity: 63.44% Conservative: 29  
 Best Local Similarity: 53.05% Mismatches: 53  
 Query Match: 39.95% Indels: 49  
 Gaps: 12

US-10-630-636-7 (1-306) x CK25717 (1-961)

```

QY 13 ALAAlaAlaAlaATPThrArgGluAspAspLysAlaPheGluAsnAlaLeuAlaAla 32
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 197 TCATCTTCAGTACGTGAGCAGAGAAAGAAAGAAAGCTTTGAGATGCGATGCTGTG 256
QY 33 CysAlaAlaProProProAlaAspGlyGlyAlaProAspAspAspTrp----- 48
Db 257 -----PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluVal 274
   49 -----PheAlaAlaLeuAlaAlaSerValProGlyAlaArgSerAlaGluVal 65
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 275 GAACAGCAATGAGGAGCACTTGTCTTATGTCCTCC--ACAAAGACCATGATGACCTT 331
QY 66 ArgArgHisTyrGluAlaLeuValGluAspValAlaAlaIleAspAlaGlyArgValPro 85
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 332 AAGGAAACATTACCAAGTACTTTTGAAAGATGTCACCACTTGAAGCAGGCCAAGTCCA 391
QY 86 LeuProArgTyrAlaGlyGluGlu---SerAlaAlaProProAspGlyAlaGlyAlaAla 104
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 392 ATACCCCAATTACAAAGAGAAAGAGCTTCTCTTCACTAAAGAAAGTTAATCTAGGA 451
QY 105 ALAAlaAlaSerLysAspGlyGlyHisArgArgAspGluArgLysGlyGlyGly 124
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 452 TATCTGGGTCAAGTACGACGGA-----CGGCGCTCCAAATTGGTTATTCAAATGGA 502
QY 125 TyrAsp-----GlyGlyLysSerCysSerLysAlaGlu 135
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 503 TTTTCAGGTACGACCCATGACCCGATTGGCGATGGGAAAGGAAATTCAGGCTCGAA 562
QY 136 GluGluArgArgLysGlyLysProTyrThrGluGluGluHisArgLysLeuLeuGly 155
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 563 CAAAGAGAGCAAGAAAGCATACCGTGGACTGAAGAAAGCAATAGCTGTTTCTAGGT 622

```



```
Oy 156 LeuAspLysPheGlyLysGlyAspTrpArgSerIleSerArgAsnPheValIleSerArg 175
Db 623 TTAGCAAAATTGGGAAAGAGATTGAGAAATTTCAGGAATTTGTTATATCTCGA 682
Oy 176 ThrProThrGlnValAlaSerHisAlaGlnLysTrpPheIleArgLeuAsnSerMetAsn 195
Db 683 ACACCAACACAAAGTAGTGTAGTCATGCCAGAAATACCTTATTCGTTGAAATTCATGAAAT 742
Oy 196 ArgAspArgArgSerSerIleHisAspIleThrSerValThrAlaGlyAspGlnVal 215
Db 743 AGAGATAGAAAGAGCTCAAGTATTCACGACATTCATCAATCATATATGAGAGACGTT 802
Oy 216 AlaAlaGlnGlnGlyAlaProIleThrGlnHisGlnAlaThrGlyAsnProAla---Ala 234
Db 803 TCAACTCATCAA---GCTCCAATTAAGGCCAACAGTGAAATCCGAATCCATCAATCCA 859
Oy 235 AlaAlaLeuGlyProProGlyMetLysHisHis-----HisHisHisHisProGly 251
Db 860 GCTGCTCTTGA---CCCTCAGTCAGACAGAACCAACCCCAATATGATATGATGAGC 916
Oy 252 GlyAlaProProMetProMetTyrSerAlaAlaProMetGlyHisProValAla 270
Db 917 -----ATGTAC---GGTGCTCCGATGGGTCACCCCCTCCCT 949
```

Search completed: July 30, 2005, 01:24:41  
Job time : 3698 secs

**This Page Blank (uspto)**

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☒ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**This Page Blank (uspto)**